DOCTOR OF PHILOSOPHY

An analysis of the utility of quantitative faecal immunochemical tests in screening and symptomatic populations

McDonald, Paula Jane

Award date:
2016

AN ANALYSIS OF THE UTILITY OF QUANTITATIVE FAECAL IMMUNOCHEMICAL TESTS IN SCREENING AND SYMPTOMATIC POPULATIONS

BY PAULA JANE MCDONALD MSc CSci FIBMS

AN ANALYSIS OF THE UTILITY OF QUANTITATIVE FAECAL IMMUNOCHEMICAL TESTS IN SCREENING AND SYMPTOMATIC POPULATIONS

By Paula Jane McDonald MSc CSci FIBMS

Submitted for the degree of Doctor of Philosophy

2016

School of Medicine, University of Dundee

CONTENTS

Tables

FIGURES

ABBREVIATIONS

A&E	Accident and Emergency
AJCC	American Joint Committee on Cancer
APC	Adenomatous Polyposis Coli
APCS	Asia-Pacific Colorectal Screening
AUC	Area Under the Curve
BMI	Body Mass Index
BoSS	Bowel Screening System
BSG	British Society of Gastroenterology
CA-125	Cancer Antigen 125
CEDAR	Cost-Effectiveness of a Decision Rule for Abdominal Complaints in Primary Care
CHI	Community Health Index
CI	Confidence Intervals
CLSI	Clinical and Laboratory Standards Institute
CPA	Clinical Pathology Accreditation (UK) Ltd
CRCSC	Colorectal Cancer Screening Committee
CT	Computed Tomography
CV	Coefficient of Variation
DCBE	Double Contrast Barium Enema
DD	Diverticular Disease
DNA	Deoxyribose Nucleic Acid
EGTM	European Group on Tumour Markers
EQAS	External Quality Assessment Scheme
EWG	European Working Group
FFLT	FIT as a First-Line Test
f-Hb	Faecal-Haemoglobin Concentration
FIT	Faecal Immunochemical Test for Haemoglobin
FOB	Faecal Occult Blood
FOBT	Faecal Occult Blood Test
FS	Flexible Sigmoidoscopy
GP	General Practitioner
gFOBT	Guaiac Faecal Occult Blood Test
Hb	Haemoglobin
HPP	Hyperplastic Polyp
HRA	Higher (High and Intermediate) Risk Adenoma
IARC	International Agency for Research into Cancer
IBD	Inflammatory Bowel Disease
IC	Interval Cancer
IDA	Iron Deficiency Anaemia
IfU	Instructions for Use
ISD	Information Services Division

IT	Information Technology
JAG	Joint Accreditation Group on GI Endoscopy
KPI	Key Performance Indicator
K-ras	Kirsten rat sarcoma
LRA	Low Risk Adenoma
M2-PK	Pyruvate Kinase Isoenzyme Type M2
MSH1 and 2	MutS Homolog1 and 2
NCSS	National Cancer Screening Service
NHS	National Health Service
NIBSC	National Institute for Biological Standards and Control
NORCCAP	NORwegian Colorectal CAncer Prevention trial
NPV	Negative Predictive Value
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
OBD	Organic Bowel Disease
p53	Tumour protein p53
PCP	Primary Care Provider
PK	Pyruvate Kinase
PLCO	Prostate, Lung, Colorectal and Ovarian
PMS1	Postmeiotic segregation increased 1
POCT	Point-of-Care Test
PPV	Positive Predictive Value
QIS	Quality Improvement Scotland
RCT	Randomised Controlled Trial
ROC	Receiver Operator Characteristic
RPHA	Reversed Passive Haemagglutinin Antibody
SBoSC	Scottish Bowel Screening Centre
SBoSL	Scottish Bowel Screening Laboratory
SBoSP	Scottish Bowel Screening Programme
SD	Standard Deviation
SEER	Surveillance, Epidemiology, and End Results
SIMD	Scottish Index of Multiple Deprivation
SNBTS	Scottish National Blood Transfusion Service
SOP	Standard Operating Procedure
STARD	STAndards for Reporting of Diagnostic accuracy studies
TNM	Tumour Nodes Metastases
UK	United Kingdom
UKNSC	United Kingdom National Screening Committee
US	United States
WEO	World Endoscopy Organization
WHO	World Health Organization

Acknowledgements

Many people have helped me complete this thesis. I would particularly like to thank Professor Robert Steele and Dr Bill Bartlett for supervising my work. I would like thank Professor Callum Fraser for his expert guidance and proof reading.

I am also grateful to friends and colleagues from within the Scottish Bowel Screening Centre for their help and support. My fellow students Gillian Libby and Jayne Digby listened to my questions and queries. Linda Brownlee, my manager, for encouraging me to attend meetings and further the research aspects of my laboratory role. I would like to thank all of the laboratory staff who took up the training and performed the new test with enthusiasm.

I would like to thank my family and friends who have supported me and encouraged me over the past six years. I couldn't have done it without you.

Declaration

I declare that I am the author of this thesis and the work of which it is a record has been completed by me. I have consulted all the references cited and any quotations taken from these references have been highlighted in quotation marks. This thesis has not been submitted, either in the same or different form, to this or any other University for a degree.

Signed _____ Date _____

Paula Jane McDonald

I certify that Paula Jane McDonald has spent the equivalent of at least nine terms on research work under my supervision and that she has fulfilled the conditions of Ordinance 39 so that she is qualified to submit this thesis for the degree of Doctor of Philosophy.

Signed _____ Date _____

Prof Robert JC Steele

Abstract

Background

It has been demonstrated, in 4 large randomised control trials (RCT), that screening for colorectal cancer (CRC) using annual or biennial guaiac faecal occult blood tests (gFOBT) reduces mortality and incidence. The faecal immunochemical test (FIT) uses technology that is analytically more sensitive and specific for human haemoglobin (Hb) than gFOBT.

Methods

An evaluation of the OC-Sensor Diana quantitative FIT analyser and prospective analysis of a single estimate of faecal haemoglobin concentration (f-Hb) in two clinical settings; the Scottish Bowel Screening Programme and patients referred from primary care to endoscopy services.

Results

Uptake, in the cohort offered screening with FIT as a first-line test, was 4.8% higher than that seen contemporaneously in the Scottish Bowel Screening Programme. This returned to pre study levels when the study ceased and gFOBT was reintroduced. The cohort offered quantitative FIT had a positivity of 2.4% compared to 2.1% in the programme overall. Clinical outcomes, during the evaluation period, in the study cohort and the screening programme were similar.

40,125 participants returned a FIT sample device and 38,720 had their f-Hb measured. An observational study of f-Hb by sex and age, using the 97.5th percentile as a potential upper reference limit, and 90% confidence intervals (CI) showed 519 ng Hb/ml buffer (90% CI: 468 – 575) for men and 283 ng

Hb/ml buffer (90% CI: 257 – 316) for women. When the data was partitioned by age quintile, f-Hb increased with age in both sexes.

Quantitative FIT and endoscopy were completed by 280 patients referred from primary care for endoscopy (median age: 63 years, range: 18 to 84 years), 59.6% were female. Six (2.1%) participants had CRC, 23 (8.2%) high-risk adenoma (HRA: >3 adenomas or any > 1 cm), 31 (11.1%) low-risk adenoma (LRA), and 26 (9.3%) inflammatory bowel disease (IBD) as the most serious diagnosis. Those with CRC had median f-Hb of > 1000 ng Hb/ml buffer. Using f-Hb with a cut-off of 50 ng Hb/ml buffer, negative predictive values of 100%, 94.4%, 93.4% and 93.9% were found for CRC, HRA, LRA and IBD.

Conclusions

The introduction of quantitative FIT into population screening and symptomatic settings has the potential to optimise referral for endoscopy for those who have evidence of small amounts of bleeding, thereby improving outcomes and reducing the current burden on endoscopy services.

Publications Arising from this Work

McDonald PJ, Strachan JA, Digby J, et al. Faecal haemoglobin concentrations by gender and age: implications for population-based screening for colorectal cancer. *Clinical Chemistry and Laboratory Medicine* 2012;50: 935-40.

McDonald PJ, Digby J, Innes C, et al. Low faecal haemoglobin concentration potentially rules out significant colorectal disease. *Colorectal Disease* 2012;15: 151-159.

Steele RJC, **McDonald PJ**, Digby J, et al. Clinical outcomes using a faecal immunochemical test for haemoglobin as a first-line test in a national programme constrained by colonoscopy capacity. *United European Gastroenterology Journal* 2013;1: 198-205.

Publications Associated with this Work

Fraser CG, Digby J, **McDonald PJ**, et al. Experience with a two-tier reflex gFOBT/FIT strategy in a national bowel screening programme. *Journal of Medical Screening* 2012;19: 8-13.

Digby J, **McDonald PJ**, Strachan JA, et al. Use of a faecal immunochemical test narrows current gaps in uptake for sex, age and deprivation in a bowel cancer screening programme. *Journal of Medical Screening* 2013;20: 80-5.

Digby J, **McDonald PJ**, Strachan JA, et al. Deprivation and faecal haemoglobin: implications for bowel cancer screening. *Journal of Medical Screening* 2014;21(2): 95-97.

Digby J, Fraser GF, Carey FA, **McDonald PJ**, et al. Faecal haemoglobin concentration is related to severity of colorectal neoplasia. *Journal of Clinical Pathology* 2013;66: 415-419.

CHAPTER 1: Detection of Colorectal Cancer (CRC)

1.1 Introduction

Colorectal cancer (CRC) continues to be the third most commonly diagnosed cancer worldwide. It is ranked as the fourth most common cause of cancer deaths for men and women. Incidence varies greatly across the world with developed countries accounting for nearly 60% of cases. The developing countries are experiencing an increase in the incidence of CRC that may be due to recent changes in environmental factors such as adoption of a western diet, increased alcohol consumption and reduction of exercise (1).

Annually, in the United Kingdom (UK), around 40,000 people are diagnosed with CRC. In 2012, there were approximately 16,000 deaths from CRC in the UK. The incidence of CRC rises sharply at age 50 years in men and 55 years in women, with the sharpest increase seen in males. Incidence rates in men rose by an average of 1% each year between 1979 and 1999 (2). It might therefore be expected that the burden of CRC will increase as the population ages.

Most cases of CRC present to primary care symptomatically or are detected during the investigation of a different condition, such as anaemia. Prognosis and outcome are related to the stage of the disease at presentation. Two staging systems are commonly used to classify CRC. The American Joint Committee on Cancer (AJCC) advocate using the size of tumour, lymph node involvement and whether there are metastases (TMN) to group bowel cancers into numerical stages 0 – I, II, III, IV. (3). The second method is Dukes' staging which uses the depth of penetration through the gut wall –

A means the CRC has only penetrated the inner gut wall through to D where the CRC has spread to elsewhere in the body.

Survival rates increase dramatically when CRC is detected early – in the symptomatic population, CRC diagnosed at Dukes' stage A have a five year survival of 93.2%, whereas Dukes' stage D CRC have a five year survival of 6.6% (4). Screening enables the detection of early disease in those who are apparently well and it has been shown that early detection through asymptomatic population screening reduces CRC mortality (5). In addition, removal of sufficient numbers of adenomas, possible precursors of CRC, also decreases disease incidence in the screened population (6).

In this Chapter, the following topics are reviewed: models of CRC carcinogenesis, use of tumour markers, patients with symptoms presenting in primary care, the principles of screening, detection methods, initiatives to increase uptake in screening programmes, some of the adverse effects of screening and novel approaches to CRC detection.

1.2 Models of CRC Carcinogenesis

We can consider carcinogenesis as a process that results in the progression of a normal gut epithelial cell to a malignant CRC cell. Genetic changes accumulate within the affected cell line that confer a selective growth advantage over others. The mechanisms underlying these transitions are complex and may involve genetic mutations or epigenetic changes affecting oncogenes, tumour suppressor genes or mutator genes (3).

Fearon and Vogelstein (7) proposed a multistep, genetic basis for the development of CRC, and the adenoma-carcinoma sequence is generally

accepted as the stepwise model for the formation of most CRC. In this model, genetic mutations that allow cell proliferation to continue unchecked are accumulated. Apoptosis is switched off and angiogenesis occurs, leading to carcinogenesis. Although this provides a model for the pathway preceding carcinogenesis, it must be remembered that approximately 40% of the western population will develop adenomas, but that only 3% of these adenomas have malignant potential (8). Key tumour suppressor genes are: APC, p53 and DNA mismatch repair genes MSH2, MLH1 and PMS1 and 2. Key oncogenes are K-ras and DNA binding protein genes myc and myb. The absence or presence of these mutations varies from tumour to tumour and may have a bearing on disease prognosis and response to therapy (9).

A *'de novo'* pathway in which CRC do not derive from pre-existing adenomas, was suggested in the 1980s. However, this model remains controversial. It was first proposed in a retrospective review of patients undergoing follow-up of polyps over six years. It was estimated that this type of CRC may be responsible for one-third of all CRC detected (10). A second retrospective study of 1630 tumours found that these *'de novo'* CRC were not a single group and were more often found in the right colon (11). When considering methods for the detection of CRC, this group must be taken into account since they may not bleed as in the adenoma-carcinoma sequence and they may be more difficult to visualise during endoscopy.

1.3 The Presence of Haemoglobin (Hb) in Faeces as a Tumour Marker

Tumour markers are usually proteins produced by normal cells in small amounts that may be over expressed in cancerous conditions. Examples of so called specific tumour markers are prostate specific antigen for prostate cancer and CA-125 for ovarian cancer. The diagnostic efficiency of those tests and other issues may however preclude their use in diagnostic or

screening contexts. Detection of the presence of haemoglobin (Hb) in faeces, often described as faecal occult blood (FOB) may be an indicator of significant colorectal neoplasia (higher risk adenoma (HRA) and/or CRC) since many lesions go through a process of vascularisation during their development. In this case we are not witnessing over expression of a marker, but unusual release of blood by the affected tissue. Early studies using red calls labelled with Cr ^{51} showed that, in 555 faecal samples from 80 patients with clinical suspicions of blood loss, there were detectable amounts of Hb in faeces measured by Cr ^{51} and guaiac methods (12). However, this early study indicated that the presence of Hb did not always indicate significant neoplasia and may be an indicator of the presence of other organic bowel disease, such as inflammatory bowel diseases (IBD) or may have no obvious underlying cause. The diagnostic efficiency for colorectal cancer is therefore reduced because of false positive results.

There is no perfect tumour marker and the benefits and disadvantages of examining a particular tumour marker must be weighed up when making a decision regarding whether or not to use it (13). The characteristics of the ideal tumour marker are shown in Table 1.

Table 1 Characteristics of a tumour marker

(1)	The marker must be produced by the tumour
(2)	Should be produced in sufficient quantities to be detected
(3)	Must have a half-life that allows monitoring to take place
(4)	Sampling must be relatively easy
(5)	Concentrations must be relative to tumour progression

Adapted from Holland-Frei Cancer Medicine (13)

Where there is blood loss into the bowel, the detection of Hb in faeces is dependent on several factors: the principle of the method used, site of blood loss, size of lesion and type of lesion. Microflora in the gut degrade Hb into

its constituent parts and these degradation products may become undetectable by any test for the presence of Hb in faeces, commonly called faecal occult blood tests (FOBT).

Increased transit time will reduce the amount of Hb and this may be a particular issue in females who generally have increased transit time compared to males (14). Another consideration is whether the lesion in the bowel is developed enough to start bleeding or is a tumour that is prone to bleeding. The amount of Hb may also be below the analytical detection limit of the test method chosen or frequency of blood loss may be such that the bleed is missed at the time of sampling. It follows that the presence of Hb may enable detection of CRC, but that it may not be particularly sensitive. This issue will be explored later in this work.

Focusing on the methods for testing for the presence of Hb in faeces, the characteristics of the two approaches most commonly used are compared in Table 2. Those are the guaiac faecal occult blood tests (gFOBT) and faecal immunochemical tests (FIT) for Hb.

Table 2 Comparison of faecal occult blood tests (FOBT)

	gFOBT	**FIT**
Target	Haem	Globin
Principle	peroxidase activity using guaiac as an indicator	Immunochemical reaction
Method	Indirect	Direct
Interference	Meat, some vegetables containing peroxidases, vitamin C	None
Detects	Bleeding from stomach, small and large bowel	Bleeding from large bowel

Adapted from Colorectal Cancer in Clinical Practice (15)

1.4 Presentation of CRC in Primary Care

Abdominal symptoms such as pain, change in bowel habit or rectal bleeding are common and while they may be used as potential indicators of CRC, they are recognised as being very non-specific. In a systematic review and meta–analysis of 47 diagnostic studies in primary care from 2010, Jellema et al. (16) found that the only single reported physical symptom that that had any significance in the diagnosis of CRC was weight loss. Other individual symptoms were not useful for predicting CRC: however, certain symptom combinations, such as rectal bleeding, abdominal pain and change in bowel habit increased the sensitivity of diagnosis of CRC. Including FIT in the algorithm increased specificity. The review concluded that, due to the heterogeneity of the studies, the results may not be directly applicable to all primary care settings and further focused studies were needed.

An early Italian study assessed the effectiveness of primary care providers (PCP) at diagnosing patients presenting with symptoms (17). This study was conducted over a period of eight weeks, 159 primary care physicians were involved and 332 symptomatic patients underwent colonoscopy or double contrast barium enema (DCBE). From this group, 280 patients were eligible for inclusion in the study. The most frequent symptoms reported were lower abdominal pain (79.6%), bloating (59.6%), constipation (47.8%) rectal bleeding (40.7%), diarrhoea (30.3%) and iron deficiency anaemia (IDA) (24.6%): 90.3% of patients complained of more than one symptom and 73.2% complained of at least three symptoms. Clinicians were more likely to refer for colonoscopy if rectal bleeding was a symptom and DCBE if abdominal pain and bloating was a symptom.

The study found that both these symptoms were not risk factors for CRC. The most common symptoms in patients with CRC were abdominal pain (73%), IDA (68%), bloating (54%), constipation (51%), rectal bleeding

(44%), weight loss (37%), diarrhoea (24%) and change in bowel habits (20%). The only variables significantly associated with a risk of CRC were age greater than 50 years, IDA and weight loss.

Iron deficiency anaemia occurs in up to 5 % of the adult population in the developed world. In 5-10% of cases CRC is an underlying cause and may also be a presenting feature in primary care. Current guidelines, written by the British Society of Gastroenterology (BSG), indicate that, where there is no history of poor diet, use of nonsteroidal anti-inflammatory drugs (NSAIDs), family history or blood donation follow up should include colonoscopy to investigate the possibility of occult blood loss due to CRC (18).

The National Institute for Health and Care and Excellence (NICE) is a special health authority set up in England who produce guidance and standards in relation to provision of health services. Within Scotland, the Scottish Intercollegiate Guidelines Network (SIGN) provide evidence based recommendations regarding referral strategies for patients presenting to primary care. Guidelines reduce variation and improve the quality of care by gathering and reviewing the most up to date evidence and presenting it in an accessible form for non-experts. Both organisations have produced guidance in relation to CRC. SIGN 126 Diagnosis and management of colorectal cancer and (19) NICE NG 12 Suspected Cancer: recognition and referral (20) recommend that patients over the age of 40 years, reporting to primary care with rectal bleeding are referred for colonoscopy, with NICE also including abdominal pain, weight loss and IDA in its list of symptoms requiring urgent referral.

Offering all patients presenting to primary care with abdominal symptoms colonoscopy is expensive and potentially hazardous. In addition a woman presenting with abdominal distension or bloating with or without abdominal

pain, feeling full quickly, or difficulty eating, may have other serious causes of their symptoms requiring investigation via other pathways (e.g. ovarian cancer). As described above, most patients presenting to primary care with non-specific symptoms do not have CRC. Controversially, the most recent revision of the NICE guideline (20) includes an update that states that FOBT can be offered to this group, this could be considered a retrograde step as it encourages reintroduction of gFOBT in assessment of symptomatic patients.

The use of a single faecal sample obtained by digital rectal examination using gFOBT in assessment of symptomatic patients has been discouraged in national guidelines (21). In fact, these tests have many serious disadvantages and should be considered obsolete as point-of-care tests since they may actually slow progression to endoscopy (22). Newer FIT and faecal calprotectin (inflammatory marker) tests may be able to replace the gFOBT in the assessment of those presenting at primary care. Analysis of 400 patient samples from the cost-effectiveness of a decision rule for abdominal complaints in primary care (CEDAR) study found that the diagnostic accuracy of faecal calprotectin and faecal haemoglobin concentration (f-Hb) analyses depended on the size of the adenomas present and was not sufficient to detect lesions <1 cm in diameter. The tests were able to rule out organic bowel disease (OBD), but would not be useful for diagnosis of significant neoplasia (23).

1.5 Screening for CRC

Wilson and Jungner stated that the aim of screening is to diagnose disease during the early stages of its pathogenesis in people who do not have symptoms. They put forward 10 principles of screening as outlined in Table 3 (24). CRC is an excellent candidate for screening; it has a well described

development, it is a significant health issue and there are interventions available that reduce morbidity and mortality.

Table 3 Principles of screening

(1)	The condition sought should be an important health problem.
(2)	There should be an accepted treatment for patients with recognised disease.
(3)	Facilities for diagnosis and treatment should be available.
(4)	There should be a recognisable latent or early symptomatic stage.
(5)	There should be a suitable test or examination.
(6)	The test should be acceptable to the population.
(7)	The natural history of the condition, including development from latent to declared disease, should be adequately understood.
(8)	There should be an agreed policy on whom to treat as patients.
(9)	The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
(10)	Case-finding should be a continuing process and not a "once and for all" project.

Adapted from WHO Report Principles and practice of screening for disease (24)

The most robust evidence for screening for CRC comes from a meta-analysis of randomised controlled trials (RCT) using triple card gFOBT followed by colonoscopy for those who had a positive test result (25). As a result the detection of Hb in faeces, including the use of FIT methods, is the approach used in many European countries (26).

This type of screening is a two-step approach, where any person who has a positive faecal test result is referred for bowel visualization. However, in a one-step process the investigation is direct referral for bowel visualisation and includes flexible sigmoidoscopy and colonoscopy (27). The use of colonoscopy as the first-line test for the detection and removal of clinically

significant lesions was endorsed by the US Preventive Services Task Force in 2008 (28).

In addition, testing may be performed on a selected population (mass screening) or efforts may be directed towards the individual. Countries in the European Union tend to adopt nationally organised mass screening whereas, in the United States (US), screening tends to be undertaken by case-finding or opportunistic interventions in primary care (29). The Survey of Colorectal Cancer Screening Practice in Healthcare Organisations and National Health Interview Survey, reported on adherence to screening guidelines among PCP (30). A total of 1175 PCP and 11365 adults > 50 years of age responded to a questionnaire about FOBT implementation and follow-up of positive test results. It was recommended that home triple card gFOBT are used for screening: however, 32.5% of doctors responded to the questionnaire stating that they use once only digital rectal examination during the consultation. Full endoscopic examination of the bowel was recommended for patients with a positive test result, however, 29.7% of doctors followed up a positive result with a further gFOBT test.

Of the adults who completed a questionnaire, one-third said they had a single test in the consultation and a further one-third said that they had no follow-up after a positive test result. One drawback to the study is that the results were self-reported and may contain inaccuracies. Poor adherence to stated guidelines can potentially reduce the effectiveness of screening and the reduction in disease specific mortality reported in the RCT may not be attained under these conditions. It is important to note that nationally organised programmes are easier to monitor through the central collection of data items which give an indication of effectiveness (31).

A comparison of US and European screening modalities concluded that the number and combination of screening modalities available could lead to variability in overall outcomes (29). However, sharing the outcomes of the different screening programmes at well organised international meetings could be a way to develop and improve future efforts (26).

Screening, post-screening surveillance and increased awareness of signs and symptoms in the general population and subsequent presentation in primary care since the inception of screening programmes all put extra pressure on endoscopy services. A review of the demand for endoscopy capacity within England, showed increasing demand from baseline symptomatic referrals, general practitioner (GP) referrals, screening and awareness campaigns when they are run (32).

Due to the recommendations that all these different groups of patients should be referred for bowel visualisation, the endoscopy requirement is increasing year on year and resources are struggling to keep up with demand. What is required is a way to triage those who are at higher risk of CRC and would benefit most from endoscopy.

1.6 Evidence Base for Use of Guaiac Faecal Occult Blood Tests (gFOBT) as a Screening Modality

Within a screening context, the sensitivity of gFOBT for CRC is difficult to estimate but, using interval cancer (IC) data (CRC that are diagnosed after a negative screening test in the interval until the next screening invitation), the most commonly used triple test (Hemoccult ® (Beckman Coulter, Fullerton, CA)) seems to detect in the region of 50% of CRC in a population that accepts screening (33). The accepted dogma is that this low clinical sensitivity is due partly to the relatively low analytical sensitivity of Hemoccult for blood and that CRC may bleed intermittently or not at all. On

the other hand, the specificity (percentage of disease free individuals with a negative test) is around 98%.

Although this seems high, the majority of the population do not have CRC: and, because 2% of a large number in a population-based CRC screening programme is a considerable cohort, this results in a high false positive rate. It should also be noted that most studies used non-rehydrated Hemoccult: rehydration increases both analytical and clinical sensitivity by lysing red cells and exposing more haem to react with the peroxidase developer. However, the increase in clinical sensitivity has an adverse effect on specificity. This is an issue with all methods using FOB as a means of detecting colorectal neoplasia.

The principles of screening indicate that CRC is a good candidate for screening. Data regarding early detection suggest that this leads to improved prognosis. Therefore, it could be concluded that detecting CRC through screening will have a positive effect for the whole population. However, the screening process is associated with biases that have the effect of making screen-detected disease appear to have a better prognosis than symptomatic disease. The terms used to describe these biases are volunteer bias (and non-responder bias), length bias and lead-time bias.

Volunteer bias results from invitations to be screened being taken up more readily by people who are health conscious. In other words, people who accept screening invitations are more likely to experience a better outcome if disease is detected; for example, they may be more likely to take exercise and less likely to smoke. Those people who do not respond to the invitation to be screened may be different to those who did accept. Length bias is a consequence of intermittent screening tests tending to pick up slow growing disease that is more likely to be associated with a good prognosis than more

aggressive, fast growing disease that is likely to present between screening episodes. Finally, lead-time bias is an inevitable consequence of early diagnosis: detecting disease at an early stage of its development inevitably leads to an observed improved duration of survival simply by virtue of shifting the point of diagnosis forward in time (15).

To prove that screening is effective, these biases must be eliminated and, to do this, population-based randomised controlled trials (RCT) are required. In these RCTs, it is essential that the group randomised to be offered screening must be analysed as a whole, including those who do not choose to participate in the process and those who are diagnosed with IC. The disease-specific mortality in this group must then be compared with that seen in a randomly selected control group which is not offered screening and, only if a significant improvement is seen in the test group, can screening be considered to be beneficial.

The original population-based RCTs of CRC screening utilised gFOBT and, since the results of these trials proved unequivocally that screening for CRC is effective, it is worth examining them in detail. Five studies were carried out, in the United States, England, Denmark, France and Sweden, and all employed the gFOBT Haemoccult (Hemoccult in US market).

The study from the United States, which took place in Minnesota, randomised volunteers into three groups, one group offered annual screening (15,930 men and women), one group offered biennial screening (15,587 men and women) and a third observation group (15,124 men and women). It should be noted that, uniquely amongst the RCT, the Hemoccult was rehydrated resulting in a 10% positivity and, as a result, in the group offered annual screening, 38% had a colonoscopy on at least one occasion. A positive test result was followed by colonoscopy and CRC mortality was

reduced by 21 % in the biennial group and by 33% in the annual group after 13 and 18 years follow up (34). In addition, a significant reduction in incidence was noted in annual and biannual testing. This has not been observed in any of the other trials of gFOBT screening and is probably the result of the high colonoscopy rate leading to a high rate of polypectomy and removal of potentially neoplastic lesions (6).

The study carried out in England took place in Nottingham and 150,000 subjects were randomised by household (35). The group offered screening were aged between 50 and 74 years and offered biennial non-rehydrated Haemoccult. Colonoscopy was offered after a positive test result. In the first (prevalence) round, the positivity was 2.0% and, in the subsequent (incidence) rounds, it fell to 1.2%. In the course of five screening rounds, 60% of the group which was offered screening completed at least one screening test. The screen-detected CRC were diagnosed at an earlier stage, 57% at Dukes' stage A, with more favourable outcomes. However, there were a large number of IC with 50% of the CRC diagnosed in those who had accepted at least one screening test not detected.

Thus, in this study, gFOBT was associated with a relatively low uptake and relatively poor sensitivity. Nevertheless, the CRC specific mortality in the group offered screening showing a significant reduction of 15% when compared to the control group after a median of 7.8 years of follow-up. At a median of 19.5 years of follow-up, the reduction in CRC mortality was maintained at 13% and, when adjusted for non-compliance, the reduction was found to be 18% (36). It should be noted, however, that despite the fact that 615 adenomas greater than 10 mm in diameter were removed from individuals in the intervention arm, no significant difference in CRC incidence could be detected.

Screening programmes may have a “halo” effect and the Nottingham study illustrates this. In the control group, the percentage of patients presenting with CRC at Dukes’ stage A increased from 9% in the first half of the recruitment period to 28% in the second (37). Thus, it seems that the very presence of the screening programme had an effect on individuals who were not invited for screening and this may be related to increased awareness of the significance of rectal bleeding. On the same theme, significantly fewer emergency admissions for suspected CRC were seen in the group offered screening when compared to the control group indicating that screening also had the effect of reducing the number of individuals presenting as emergencies (38).

In the Danish trial, in Funen, 61,933 subjects were randomised into a group that was offered screening by means of biennial Haemoccult testing or into a control group. The uptake in the first round was 67% and more than 90% accepted repeated screening invitations, presumably owing to the fact that those who did not accept the first invitation were excluded. As in the Nottingham study, screen detected CRC were picked up at an earlier stage with 48% at Dukes’ stage A and IC accounted for 30% of all CRC diagnosed in those offered screening at least once. Thus, screening in the Danish Study performed in a very similar way to that of the Nottingham Study and the disease specific mortality reduction after five years was 18% (39) (40).

In the French study, in Burgundy, a randomised approach was not adopted (41). Rather, small geographical areas were identified and these were either offered screening or acted as controls. Again, a non-rehydrated Haemoccult gFOBT was used on a biennial basis and offered to individuals between the ages of 50 and 74 years. A total of 91,199 were offered screening. The positivity was 1.2% in the first round but increased slightly in further rounds and uptake in the first round was 52.8% remaining fairly constant thereafter.

Reflecting the results of the Nottingham and Funen studies, the mortality reduction seen in those offered screening was 16%.

In Gothenburg, Sweden, 68,308 citizens aged 60 to 64 years were randomised into either a control group or a group that was offered biennial Hemoccult II (42). Uptake was 70% and those with a positive test result were investigated by means of flexible sigmoidoscopy (FS) and a DCBE. After a mean follow-up period of nine years from the last screening episode, a statistically significant reduction in CRC mortality of 16% was seen.

These five studies are of the utmost importance since they are the only RCT of population screening for CRC using gFOBT. These studies are usefully compared in a systematic review (43). The consistent reduction in CRC specific mortality indicates that early detection of CRC is truly beneficial. In a meta-analysis of population screening by gFOBT, it was estimated to reduce CRC mortality by 16%, increasing to 23% when adjusted for uptake. These results demonstrate that the principle of screening for CRC is sound and, although gFOBT may be considered a sub-optimal test, these results can be used as benchmarks for newer tests (27).

The United Kingdom National Screening Committee (UKNSC) advised that a demonstration pilot of biennial gFOBT screening was required to determine whether or not the results of the randomised trials could be reproduced within the UK National Health Services (44). This pilot was successfully carried out in two areas of the UK, one in Scotland and one in England (45). As a result, the UK Health Departments have now rolled out CRC screening programmes and the initial outcomes indicate that the national programmes should provide outcomes as predicted from the RCT (46) (47).

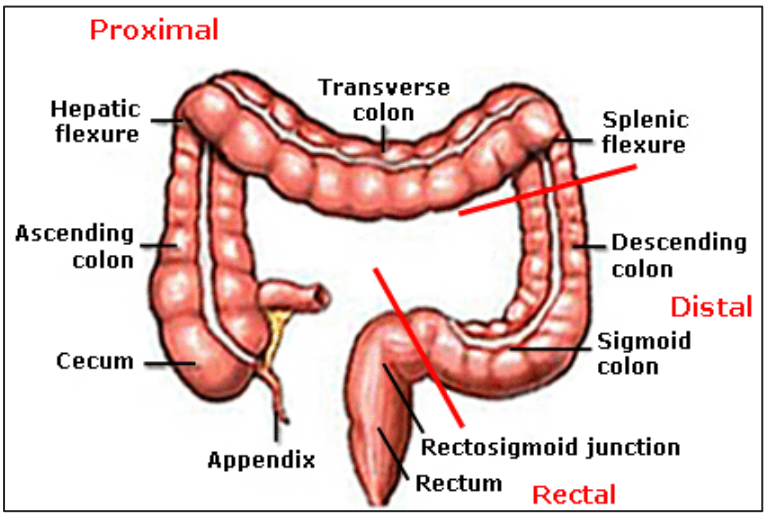
1.7 Flexible Sigmoidoscopy (FS) as a Screening Modality

Bowel visualisation is the other predominant type of screening modality. FS employs a fibre-optic endoscope to visualise the distal third of the bowel and so is able to detect lesions in the left-sided portion of the large bowel. Biopsies of suspicious looking tissue can be carried out and small polyps may be removed. Approximately two-thirds of CRC are found here. In addition, the usual practice is to refer for colonoscopy when index lesions have been found on FS. The criteria for index or high risk lesions vary but the BSG advise: 1 cm or larger in diameter, three or more adenomas, tubovillous or villous histology, severe dysplasia or malignant disease, or 20 or more hyperplastic polyps above the rectum (48).

During the FS, no sedation is required and the procedure is commonly performed by a nurse endoscopist. The bowel preparation is often a single dose, liquid enema containing sodium phosphate as this is well tolerated by the majority of people and can be performed at home or at the clinic. The FS procedure usually takes less than 20 minutes, making it less onerous than full colonoscopy.

The incidence of CRC starts to increase steeply at about 50 years. Thus, FS, or colonoscopy, performed between the ages of 50 and 60 years could be a valuable screening tool. Figure 1 shows the structure of the large bowel, this consists of the caecum, colon and rectum.

Figure 1 Diagram of the large bowel



Source - Cancer Care Ontario (49).

Holme et al report findings on once only FS or FS with FOBT compared to no screening from the NORwegian Colorectal CAncer Prevention trial (NORCCAP) (50). The cohort included 98,792 participants from three geographic areas, 78,220 in the control group and 20572 in the screening groups (split half and half between FS only and FS and FOBT). After a median of 10.9 years follow up, CRC mortality reduced by 27% and incidence by 20%. Reductions were similar in younger, 50 – 54 years old and older, 55 – 64 year age groups. The addition of FOBT led to a reduction in adherence compared with the FS group and so could be considered detrimental. Atkin et al reported on the longest running UK study (51). This was based on a multicentre trial carried out in 14 centres with 170,432 volunteers who agreed to take part in the trial. In this, the intervention group had a reduction in mortality of 31% and a reduction in incidence of CRC of 23%. However, the group offered FS were essentially volunteers and this may have led to a bias in this study.

The Italian SCORE trial, has reported on more than 10 years follow up. A total of 34,272 participants were included in the study and a non-significant reduction in mortality of 22% and a reduction in incidence of 18% were found

(52). The FS arm of the Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial reported on 77,465 subjects in the intervention group, finding that a high proportion of subjects (83.5%) accepted the initial invitation to have a screening FS. Diagnostic follow-up for those with index lesions varied according to local guidelines, but overall CRC and adenoma detection rates met expectations (53). These are significant results and offer positive insight into the effectiveness of offering endoscopy as a screening test to people in the UK.

FS may be an attractive alternative to repeat faecal testing since a single examination may be offered and other recommendations suggest an interval of five years between FS examinations. However, in a study of patients with a negative FS who had a follow up within three years 0.8% had an advanced adenoma or cancer detected (54). The protective effect of FS is dependent on the quality of a number of factors including training, bowel preparation and depth of insertion of the sigmoidoscope. The guidelines on which pathological findings trigger a colonoscopy will also affect the overall outcome: many expert panels recommend any adenoma as the trigger for colonoscopy.

Considering the advantages and drawbacks, there is great interest in screening with FS since it has the potential to significantly reduce the incidence of CRC, as well as mortality, in those who undertake the procedure through removal of polyps from the distal colon. Although uptake is not as high as for gFOBT it can be a quick procedure where bowel preparation is performed at home on the day of examination. Pain and complications are consistently reported as lower than for colonoscopy. However, this technique does not visualise the entire large bowel and the protective effect is dependent on the quality of the initial examination.

1.8 Is Colonoscopy the Reference Standard?

Colonoscopy involves manipulation of a flexible endoscope around the transverse and ascending colon to reach the caecum – reaching this point being indicative of a 'complete examination'. The bowel is prepared by modifying the diet up to two days before the procedure and administration of oral bowel cleansing preparation the day before. Patients are often offered sedation, the colonoscopy usually takes 30 – 60 minutes and must be performed by an experienced endoscopist. The bowel is inflated during the procedure and it may take 1 – 2 days for the effects of the insufflation to disappear. Colonoscopy offers a one stage screening solution, but, the procedure is not generally as acceptable to invitees as screening with gFOBT (55).

Colonoscopy is widely accepted as the reference standard in screening for CRC and is the preferred method for screening in the United States. One of the benefits to offering an endoscopic procedure as a screening test is the screening interval recommended for colonoscopy is usually 10 years and for FS 5-10 years, whereas for many non-invasive tests annual or biennial testing is advised (56).

In this setting participants who undergo colonoscopy can also benefit from prophylactic removal of polyps from the whole bowel. Removal of adenomas should reduce incidence since, according to the adenoma–carcinoma model, the presence of CRC adenoma is a predictor of risk. Analysis of data from the National Polyp Study and Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute in the US reported a lower than expected incidence of colorectal cancer in those individuals who had had adenomas removed. In this programme, adenoma removal was performed at a level that was beneficial (57).

However, there are some issues. A study conducted in Southern Germany, from 2003 – 2007, comparing 78 IC and 433 CRC detected at screening, found that, although the risk of missing CRC at colonoscopy was low, a substantial number of cancers had been missed at colonoscopy Characteristics associated with IC were positive FOBT results and incomplete colonoscopy, particularly in women. This outcome suggests that IC could be prevented by improved quality of the colonoscopies (58).

In the polypectomy study outlined above (57) three reference groups were followed up. These were from Mayo Clinic in Rochester, Minnesota, St. Mark's Hospital in London and the SEER database. The authors of the study acknowledge that generalisability could have been improved by performing a RCT and having a control group where the adenomas were left *in situ* and compared to those groups were polyps were removed, but recognition of the adenoma carcinoma sequence renders this approach unethical.

There is a slightly higher risk of perforating the bowel during colonoscopy compared with FS, patients may require sedation, air insufflation is required and the acceptability of this procedure in screening populations is low (56). Colonoscopy is an expensive procedure to perform and developing capacity for screening is costly.

As discussed earlier, the choice of screening modality effects outcomes. Table 4 shows indicative weightings for accuracy, acceptability, clinical risk and cost for gFOBT, FIT, FS and colonoscopy.

Table 4 Comparison of methods used to detect colorectal cancer (CRC)

Investigation	Accuracy	Acceptability	Risk	Cost
gFOBT	+	++	-	+
FIT	++	+++	-	++
FS	++	++	+	+++
Colonoscopy	+++	+	+++	+++

(+ LOW, ++ MODERATE, +++HIGH)

Adapted from Colorectal Cancer in Clinical Practice (15)

1.9 Can Imaging Techniques be used in CRC Detection?

DCBE has little evidence to support its use as a screening test: however, computed tomography (CT), sometimes called CT colonography, may have the potential to be used in CRC screening. This is a low risk procedure, although it uses ionising radiation, which offers the opportunity to visualise clinically significant lesions. New methods to do this without bowel preparation and using three dimensional technologies are being developed and, if proven to work effectively, such techniques could have a huge effect on the utility of CT, in this context. Sensitivity for lesions above 10 mm is > 90%: however, it is poor at detecting lesions smaller than this (59).

Bowel preparation for CT may be cathartic, or non-cathartic using radio-labelled faecal tagging. It has been demonstrated that invitation for screening with non-cathartic CT colonography has a significantly increased uptake compared to colonoscopy: however, the colonoscopy identified significantly more advanced neoplasia per 100 participants than CT colonography (60). Where cathartic bowel cleansing is used, this may offer the opportunity to equal the detection rates of colonoscopy for advanced neoplasia (61).

Although the procedure is low risk, relatively inexpensive and can identify polyps and other anatomical abnormalities, the main drawback to imaging techniques is that, if lesions or polyps are seen, the patient must then be referred for an endoscopic procedure or surgery for their removal. These techniques for bowel visualisation have been included in the recommendations of the US Preventive Services Task Force and the American Cancer Society as alternative screening tools, however they have not been widely adopted for use in population screening for CRC in other countries (29).

1.10 Maximising Uptake

The effectiveness of screening in detecting significant disease is a function of the sensitivity of the test and participation rates. There are issues regarding the perception of harm that may create barriers to participation in screening. In a survey of attitudes to screening modalities - colonoscopy, DCBE, FS and gFOBT - participants believed that colonoscopy was significantly more dangerous, embarrassing and inconvenient than gFOBT. It seems that, although colonoscopy might offer the best returns in CRC detection, it does not have the highest uptake of CRC screening investigations and so may not be the best strategy for population screening (55).

However, a recent study in Italy offering FS to screening participants who did not respond to an invitation to screen with FIT reported an improvement in the overall uptake of screening. In this case, people did respond to the more invasive procedure. Offering a choice of invasive or non-invasive test as a first-line test might thus improve uptake in national screening programmes (62).

Methods used to improve uptake in programmes offering gFOBT screening through the postal service include sending a pre-notification letter to the participant advising that they are about to be invited to screening. This intervention consistently shows increases in uptake over and above the standard invitation to screen (63). Sending invitations with additional messages about risk and with advocacy messages from previous participants have been trialled in Australia and are also recommended to improve participation (64).

There is evidence to suggest that participants who respond once will continue to respond to the invitation to screen. In the Danish RCT only participants who completed the first round of screening were invited again. Of the 67% who completed the first screening round 90% accepted further screenings. This would suggest that people who do accept screening are motivated to continue screening (39). Similarly, in the three pilot screening rounds performed in Scotland with gFOBT, initial uptake was 53%. Where those invited had not responded and were reinvited in the incident screening rounds, a further 15% and 12% responded and completed screening. It was also found that not responding to one round of screening did not infer that the person would decline all further invitations. The recommendation from this work was that reminders and invitations should be made to all those non responders identified, since this method can increase uptake (65) (66).

1.11 Adverse Effects of Screening

Although the primary aim of screening is to detect all CRC and significant adenomas, there are occasions where those who have participated in screening are diagnosed with CRC at a colonoscopy done outside the screening programme. In an annual/biennial gFOBT algorithm, IC is defined as those diagnosed after a negative gFOBT result, before they have the opportunity the next invitation to be screened.

The Nottingham study reported on IC in 1999 (67). CRC detected in the screened group and those not screened were investigated to determine whether there is a 'certificate of health effect' when a participant has participated in screening. It is argued that being given a negative screening test result may discourage a participant from seeking medical attention for symptoms of CRC between rounds of screening because they feel that they have already done a test and this leads them to believe that they will not have CRC. They found that rather than less favourable outcomes, the group who took part in screening and were given a negative result but consequently had an IC diagnosed had more stage A CRC and, 13 years after colonoscopy, had better survival than those who had not been offered screening.

In the first publication of IC data from the three pilot rounds of screening in Scotland, IC were defined as CRC diagnosed within two years of a negative gFOBT. IC rates were substantial: 31.2% in the first round, 47.7% in the second round and 58.9% in the third round. Although the percentages increased from round one to round three, this was due to a decrease in the numbers of screen-detected CRC rather than an increase in number of cases of IC. Examination of the clinical characteristics showed that the IC were associated with a better prognosis than CRC arising in a non-screened population and gFOBT appeared to preferentially detect CRC in men and the left side of the colon over CRC in women and in the right colon and rectum (33).

Research in Germany, where gFOBT testing has been offered since 1977 and colonoscopy has been offered through health insurance since 2002, indicates that more IC occur in women than men and that women are more likely to have an incomplete colonoscopy. It was also found that, in men, the finding of an IC was more likely to be associated with a positive gFOBT

result prior to colonoscopy. When the stage of CRC was reviewed, it was more likely to be more advanced than a CRC detected at screening colonoscopy. It was concluded, that the IC appeared to be more likely to be missed at colonoscopy than to be *de novo* after colonoscopy (58).

The number of IC in people who have participated in screening are a cause for concern. It is also concerning that IC occur more often in women. Indeed, is it ethical to promote screening for CRC where almost half the CRC in a population are missed? This provides a good example of the tension between the principles of population screening and the health of the individual patient. Repeated screening with gFOBT can detect significant neoplasia and reductions in CRC mortality have been reported (5), but it is also true that some CRC that are not amenable to detection by gFOBT and may be better detected by other approaches. These should also be included in the repertoire of tests used to detect CRC.

One way to decrease the number of IC would be to decrease the threshold for a positive result and refer more participants for colonoscopy. However, this presents the issue of over-diagnosis and increasing the exposure of healthy individuals to a risky procedure. In the Minnesota RCT, the positivity between the different variables of the study; rehydration and no rehydration of gFOBT cards were 9.8% and 2.4% respectively, however the PPVs for CRC were 2.2% and 5.6% (34).

The issue of over-diagnosis has been documented in breast screening and outcome data have been reviewed in order to measure benefit and over diagnosis (68). It is possible that certain individuals with screen detected CRC would have died of other disease before their tumour became apparent. In the retrospective analysis of the Nottingham study 30 day mortality for those with screen detected CRC was investigated. Five patients

died four days postoperatively, one of whom had had an anastomotic leak. Of those dying after 30 days but within two years, only four had Dukes' A disease, three of whom died of other malignancies. It would seem that, in this study, the scale of over diagnosis was small (67).

1.12 Novel Approaches to CRC Detection

Newer tests are emerging that may offer similar reductions in mortality and incidence to gFOBT. These tests involve analysis of DNA and newer tumour markers and may be undertaken on faeces, urine, blood or other tissue or fluid samples.

Where a patient has been diagnosed with CRC, mutations present in the tumour may be excreted into other compartments. A small study looking at quantification of somatic mutations related to CRC in 25 patients demonstrated that purified DNA from faeces and plasma could detect 95% and 50% of the mutations detected in the tumour DNA respectively and that faeces was a superior matrix to extracted plasma in terms of detection of CRC at early stages. This superiority may be due to the higher frequency of mutant DNA than normal DNA in faeces: however, the extraction of relevant DNA is more difficult due to the background noise from bacterial DNA (69).

The major benefit of finding a useful measurand in blood is that the sample can be taken immediately, in any clinical setting, and there may be more opportunity to take such samples during other visits to a healthcare provider. A comprehensive review of over 93 studies describing 70 different tumour markers detected in blood samples has been published (70). Aberrantly methylated genes are useful markers since they are cancer specific, occur early in cancer development, have an amplifiable signal and can be assayed with high accuracy. Case control studies have found the presence of

mSEPT9 to be highly correlated with colorectal cancer suggesting that it is a suitable candidate for developing a biomarker assay (71). Subsequently this assay was prospectively assessed in a cohort of asymptomatic individuals, > 50 years old and attending screening colonoscopy in Germany. The sensitivity for CRC was 48.2% and specificity was 91.5% (72).

As mentioned earlier in this Chapter, a number of steps contribute to the genesis of CRC, other areas of investigation are mutations in the APC gene, the p53 tumour-suppressor gene, the K-ras oncogene and genes involved in DNA-mismatch repair. Imperiale et al conducted a large study using a faecal DNA panel which consisted of 21 mutations: 3 in the K-ras gene, 10 in the APC gene, and 8 in the p53 gene, also the microsatellite instability marker BAT-26 and a marker of long DNA (73). This panel was able to detect 16 of 31 invasive CRC (TNM stages I, II, or III), giving a sensitivity of 51.6 %; Hemoccult detected 4 of 31 CRC, giving a sensitivity of 12.9 %. The panel detected 22 lesions that were missed by Hemoccult, whereas Hemoccult detected three lesions missed by the panel. This difference in test results was significant ($p < 0.001$). Among the 40 subjects who had adenomas with high-grade dysplasia, the faecal DNA panel detected 13 of the adenomas (32.5 %), whereas Hemoccult detected 6 (15.0 %). For the detection of other advanced adenomas (villous polyps and tubular adenomas 1 cm in diameter or larger) and for minor polyps, the sensitivities of both tests was consistently less than 20%. The sensitivity of the faecal DNA panel was four times that of Hemoccult for invasive CRC and more than twice as sensitive for adenomas. This increase in sensitivity was achieved without a loss of specificity among persons with no polyps on colonoscopy.

This work has been developed further to suggest that combining the multi-target faecal DNA test with FIT result would increase sensitivity while

keeping the numbers of patients referred for colonoscopy to a manageable level (74). In August 2014 the Food and Drug Administration approved the 'Cologuard' test for non-invasive screening of adults of both sexes, over the age of 50 years with average risk of CRC. The test is available in all states in the US, Medicare and a growing number private healthcare providers include it in their billing: however, there exists no evidence of its performance in an organised screening programme or evidence of what intervals between screening tests should be used.

A major challenge in the development of DNA tests for the detection of CRC is the selection of relevant markers. However, there is an opportunity to develop panels of markers related to all aerodigestive CRC, including those of the lungs, oesophagus, stomach, pancreas and bile duct. Clinical investigation could be directed to certain anatomical sites based on positive panel outcomes (75).

Tumours leak mutated DNA into the lumen of the bowel. They also leak other proteins that could be used as tumour markers. Faecal calprotectin is released from neutrophils in the colon when there is inflammation present. A recent study evaluated the usefulness of calprotectin and FIT in patients presenting in primary care. Where a patient has a positive test result for calprotectin or FIT, or both, and all sizes of adenomas were considered, the test did not detect Organic Bowel Disease (OBD) very well. However, when adenomas >1 cm were considered, the tests were very good at ruling out OBD. These data concern the application of calprotectin and FIT in primary and secondary care, but it could be hypothesised that any utility in these settings could extend to screening programmes (23).

Lactoferrin is a member of the transferrin family of iron binding proteins. It is active in fighting bacterial infections. Like calprotectin, elevated

concentrations can be useful in discerning whether bowel disease is organic or non-organic. Enzymes also have the potential to detect CRC. Pyruvate kinase (PK) has a regulatory role in glycolytic metabolism and the pyruvate kinase isoenzyme type M2 (M2-PK) is preferentially expressed by tumours. This makes it an ideal candidate for measuring tumour formation. A review of 10 studies assessing M2-PK showed that this could achieve an overall sensitivity of between 68.8% and 91% and an overall specificity of 71.9% and 100% (76).

Investigation of new approaches to screening for CRC with non-invasive tests is promising in terms of sensitivity and specificity for CRC: however, these recent developments do not have the same programmatic and longitudinal data that supports the gFOBT as a screening modality.

1.13 Conclusions

CRC is an important cause of morbidity and mortality. There is a significant evidence base for the primary role of screening for CRC as it provides an opportunity to detect disease at an earlier stage. Individuals have a poorer prognosis if disease is detected at a later stage and where there have been metastases to other organs. Use of gFOBT as a screening modality has a sound evidence base from RCT performed in the 1990s in which significant reductions in disease specific mortality were demonstrated.

The RCT outlined in this Chapter all used gFOBT as the screening test. However, there were differences in study protocols such as screening interval, (annual or biennial), sensitivity (testing dry or rehydrated gFOBT) and whether dietary restriction preceded sample collection. In the Minnesota study, the group offered annual screening with a rehydrated test card and undertook dietary restriction. This study had the greatest positivity

and referrals for colonoscopy and the most involved interactions for participants. However, this study reported a 33% decline in mortality at 13 years (33).

The initial mortality reductions seen in the Nottingham gFOBT study of 1996 have been followed up for 20 years and found to have been sustained (35). With each subsequent round of screening, the number of CRC in the screened population is reduced and a cumulative protective effect is seen (46). As mentioned earlier, there is also the potential to reduce incidence of CRC when repeated gFOBT is used as a screening modality as polyps are removed from the screened population. A continuing reduction in incidence of CRC has been shown at the 18 and 30 year follow-up of participants in the Minnesota trial (6) (77).

Screening with colonoscopy as the first-line test has high sensitivity and specificity for CRC. However, endoscopic examination of patients is expensive, has some risks and not all individuals wish to undertake this procedure. This means that improvements to other non-invasive tests should continue to be pursued alongside those in endoscopy.

Individuals who present to primary care often have later stage CRC than those detected by screening: gFOBT are no longer recommended for use in primary care, and current guidelines recommend urgent referral for bowel visualisation for abdominal symptoms, rectal bleeding and IDA of unknown cause (19). However, the burden on endoscopy services is increasing as more people in primary care are referred for non-specific abdominal symptoms, as well as those referred for colonoscopy through participating in a screening programme and also surveillance. This is a significant burden on endoscopy departments. The development of a test that is more sensitive for CRC in all these groups is therefore a priority.

There is no one test or algorithm to detect CRC that can be recommended for all countries and clinical settings. Each health provider must consider the benefits and drawbacks of available detection modalities and devise a strategy that serves the needs of the population most appropriately while considering the needs of the individual.

CHAPTER 2: Why Change from gFOBT to Faecal Immunochemical Tests (FIT) for Hb Detection?

2.1 Introduction

As discussed in the previous Chapter, there is sound evidence that using gFOBT in screening programmes can identify those who would benefit from bowel visualisation to detect CRC. The reference gFOBT method for these studies was the Hemoccult test. However, since the introduction of the guaiac method for detection of Hb in faeces, other methods of detection have been developed that utilise the antibody-antigen reaction. These determine the presence of human globin, a component of Hb, in faeces and are referred to as immunochemical faecal occult blood tests (iFOBT) or more correctly as Faecal Immunochemical Tests (FIT) for Hb.

A number of approaches have been taken to delivery of FIT methods. One of the earliest used on a large scale was the reversed passive haemagglutinin antibody reaction (RPHA). This method relies on the reaction between chicken erythrocytes coated with rabbit anti-Hb A and Hb in the sample (78). The Magstream® (Fujirebio, Tokyo, Japan) represents another approach using magnetic particle agglutination to measure Hb (79). However, in more common use now are lateral flow immunoassay cassettes. This particular technology is widely used in pregnancy test kits and detection of drugs of abuse.

In order to understand why it would be beneficial to move from the gFOBT to the FIT modality an understanding of the principles underlying the methods is necessary. Focussing on the three methods to be used within the work described in this thesis, namely; gFOBT, lateral flow immunoassay

and latex immunoturbidimetry assay (80) the advantages and drawbacks that each method brings to the detection of Hb in faeces will be explored.

2.2 Principle of gFOBT and FIT

FOBT have been defined as ‘Tests for faecal occult blood, detect blood in the stool that is not visible on gross inspection, usually less than 50 mg of Hb per gram of stool’ (81). This definition is obsolete with the development of newer, more analytically sensitive tests as we shall see in this chapter.

The gFOBT method has been the first line method of choice for the current Scottish Bowel Screening service. gFOBT rely on a pseudo-peroxidase mediated chemical reaction as shown in Figure 2. Guaiac resin extracted from the Guaiacum officinale tree is used to manufacture alpha guaiaconic acid. In the presence of haem and oxygen, the alpha guaiaconic acid is converted to a quinone compound - it is the pseudo-peroxidase activity of haem that enables the reaction. The addition of hydrogen peroxide to haem and alpha guaiaconic acid increases oxygen availability, speeding up the reaction and, due to the concentration of guaiac, produces a visible blue colour. The appearance of blue colour is considered a positive test result. The blue colour has a very short half-life and results must be read within 30 – 60 seconds following the addition of hydrogen peroxide (82).

Figure 2 Guaiac reaction

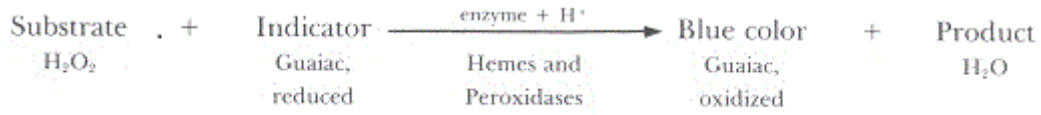


Image from Ostrow JD, Clinical Methods (81)

Commercially produced gFOBT cards consist of a cardboard housing around filter paper that is impregnated with sufficient alpha guaiaconic acid to give the manufacturer set analytical detection limit. Also included are

negative and positive control materials - the positive must turn blue when the hydrogen peroxide is added to validate the test result (82). An example of the test card and a positive test result can be seen in Figure 3.

Figure 2 Single slide guaiac faecal occult blood test (gFOBT) and positive test result

Image courtesy of the Scottish Bowel Screening Laboratory (2014)

There are many commercial suppliers of gFOBT cards and therefore many opportunities when procuring the gFOBT product for a national screening programme. The analytical detection limit of the hema-screen gFOBT (Immunostics, Ocean, New Jersey, USA) used in the Scottish Bowel Screening Programme (SBoSP) is 600 µg Hb/g faeces. This test format was developed over time to align with early work that advocated sampling three consecutive faecal samples to increase sensitivity (83). The requirement for two samples of faeces from each of three faeces in screening is the same regimen used in the algorithms in the RCT described in Chapter 1 (43).

A review of the presence of FIT in the worldwide market prepared by the Expert Working Group (EWG) on FIT for Screening of the Colorectal Cancer Screening Committee (CRCSC) of the World Endoscopy Organisation

(WEO) described 39 suppliers of qualitative FIT and 13 quantitative FIT platforms (84). The most common qualitative method used was lateral flow immunoassay. This method consists of an immunochromatographic assay strip housed inside small plastic cassette, as seen in Figure 4 (85).

Figure 4 Lateral flow immunoassay

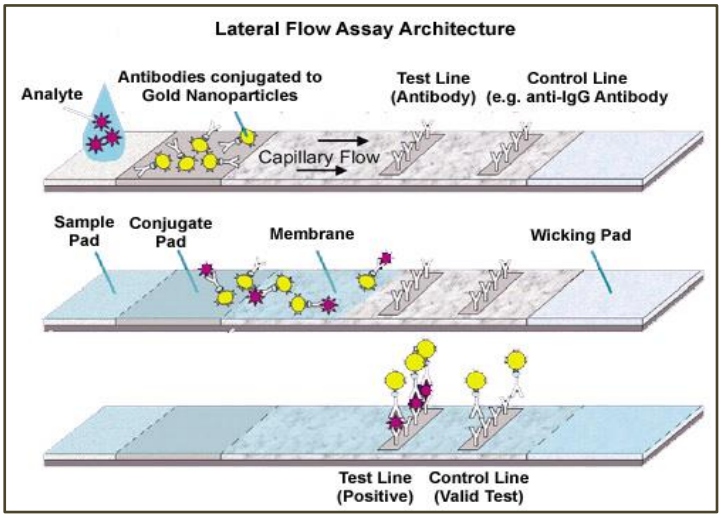


Image from Cytodiagnostics product information (85).

Faecal samples are collected in a DEVEL-A-TAB card (Immunostics, Ocean, New Jersey, USA). This is used to deliver a controlled amount of faeces into a buffer tube. The faecal suspension is delivered into the test cassette directly from the buffer tube. The test strip within the cassette has three different zones. When the liquid phase, containing the faecal sample (potentially containing human globin from Hb), is added to the well at the left hand side of the strip, it is pulled through the zones by capillary action. Primary antibodies are located in the first zone. These are mouse monoclonal antibodies directed to human globin and have gold nanoparticles are conjugated to them. Where there is human globin present in the sample an antibody antigen complex is formed here.

The next zone on the test strip has the capture antibody which is usually a polyclonal goat or sheep anti-human globin antibody. In a positive test the antibody antigen complex binds here and the gold conjugate is concentrated in this zone and a pink line appears. The final zone is the control zone which contains anti mouse immunoglobulin. This binds any free primary antibody gold conjugate. A pink line must appear in this zone to indicate the test strip is working correctly. Results are read five minutes after the addition of the faecal suspension (85). An example of positive and negative test result are shown in Figure 5.

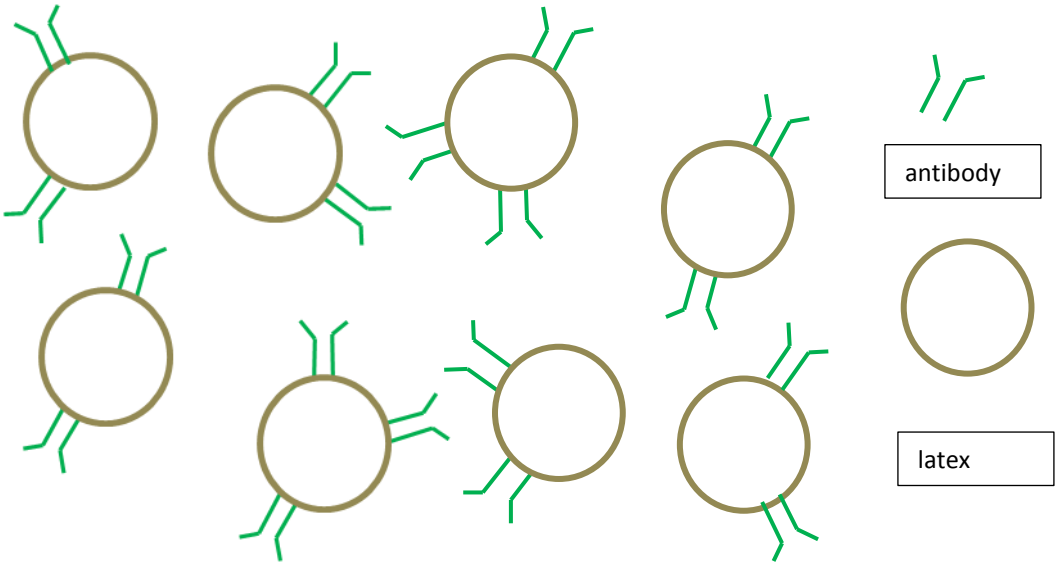
Figure 5 Positive and negative qualitative faecal immunochemical test (FIT) results

Image courtesy of the Scottish Bowel Screening Laboratory (2014)

The DEVEL-A-TAB collection device and hema-screen SPECIFIC cassette ((Fig 4.) Immunostics, Ocean, New Jersey, USA) are currently used as the second line test in the SBoSP. This system is quoted by the manufacturer to have a detection limit of 50 ng Hb/ml buffer (10 µg Hb/g faeces).

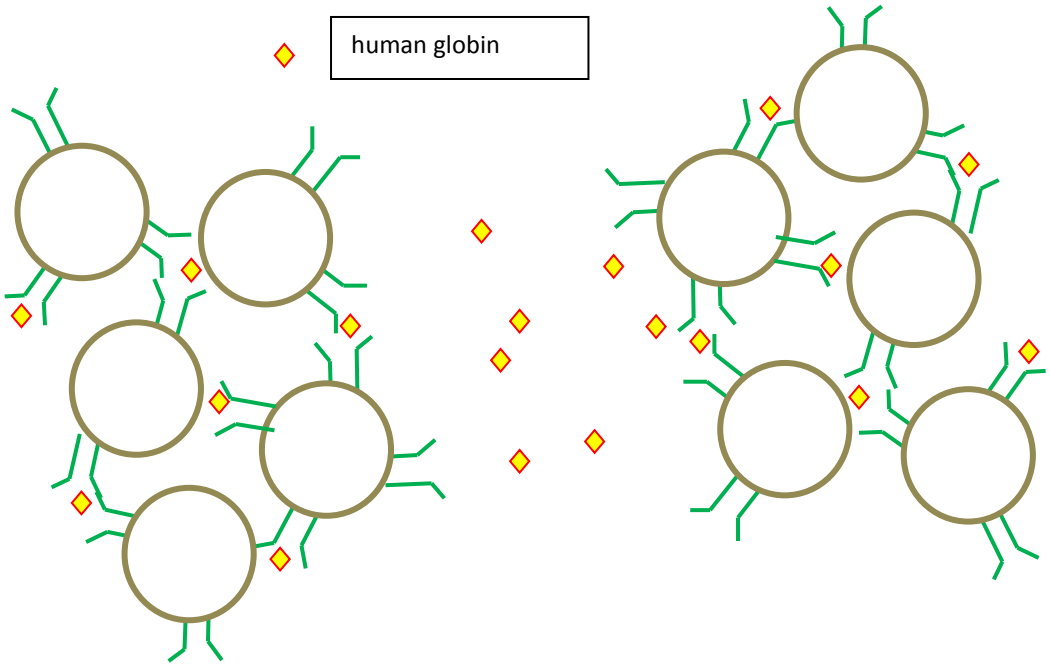
The OC Diana works on the basis of latex agglutination immuno-turbidimetry. The reagent is a latex polystyrene particle using anti HbA IgG polyclonal rabbit antibodies. (Figure 6).

Figure 6 Latex particles have antibodies directed toward globin moiety of human Hb



The instrument measures the latex agglutination in the presence of human Hb in the sample. This is shown in Figure 7.

Figure 7 Latex particles come into contact with globin in the faecal sample causing agglutination



When human Hb is present, particle agglutination occurs thus causing turbidity. The variation in this turbidity is measured photometrically at 660nm and considered as absorbed energy, is directly proportional to the Hb concentration. The immunoturbidimetric FIT method has been developed by manufacturers in Japan, where automated quantitative analysers have been used in their population based screening programmes for many years (86).

2.3 Constraints of gFOBT

The gFOBT tests have a proven track record in CRC screening programmes and have been demonstrated to be affordable. However, the methodology is problematic in that it has characteristics that will impact on the screening efficiency. Table 5 highlights potential for both false positive and false negative test results using this technology.

Table 5 Factors affecting the results of gFOBT

False-positive results.

1.	Non-human haemoglobins: red meat (myoglobin).
2.	Intake of foods that contain peroxidase activity (uncooked fruits and vegetables).
3.	Drugs: Topical iodine, aspirin, NSAIDs.
4.	Rehydration of the test kit.

False-negative results.

1.	Vitamin C (ascorbic acid) intake.
2.	Storage of cards: sunlight and heat give false positive test results.
3.	Improper sampling or development.
4.	Lesion not bleeding at time of faecal collection.
5.	Haem degradation by colonic bacteria.

Adapted from Beg (87).

A high number of false positive test results would lead to unnecessary colonoscopies being carried out and healthy individuals being exposed to unnecessary risk. A test failure, or false negative means that a case will be missed. The main considerations are discussed below:

Dietary components may cause false positive results

False positive test results may be found with excessive amounts of non-human haem found in animal proteins if they are consumed in large amounts (88). These effects may be related to regional diet, in particular high levels of black pudding consumption may cause false positive test results (89). Some raw vegetables, such as dark green leafy vegetables, horseradish and turnips, have high levels of peroxidase activity and consumption of a diet rich these foodstuffs may cause false positive test results.

A study carried out in Australia examined this effect on gFOBT and performing the development of cards at 24, 48 and 72 hours. They found that the plant peroxidase interference decreased as the sample dried but did not decrease in rehydrated tests. They concluded that, given sufficient drying time after the application of the faecal sample, (48 hours) there is no requirement for dietary restriction in screening programmes using un-rehydrated gFOBT, but that this would not work where the test is rehydrated before analysis (90).

When gFOBT were introduced, it was recommended that dietary restrictions were observed prior to the sampling of faeces. These included measures such as advising that, up to three days before the collection of samples the participant must not consume red meat because of the presence of haem compounds and that no dark green vegetables, cucumber, cauliflower or

horseradish were to be consumed to avoid artefactually increased peroxidase activity.

However, asking a participant undertaking screening to follow a strict dietary regime may adversely affect compliance (91). Sinatra et al. demonstrated that drying of the sample card prior to development reduces this interference (90) and the way many screening programmes are set up, with invitation sent through the mail and samples returned to a large central laboratory for analysis, again via the mail system this recommendation almost fulfils itself by the nature of the infrastructure of the programme.

A meta-analysis and systematic review of dietary restrictions have been conducted, Pignone et al. (92) analysed four RCTs and Konrad et al. (93) reviewed 10 case series (where the effects of challenge or restriction diets were reported), five non randomised cohort studies and five RCTs. The meta-analysis found that moderate dietary restriction would not adversely impact uptake. Both studies suggest it is reasonable to disregard previous advice to perform dietary restriction and that outcomes in large nationally organised screening programmes would not be affected. Each programme provider must look at their own situation when deciding whether to implement dietary restriction (94).

Medication can elevate positivity

Anticoagulants, such as warfarin and heparin, can cause increased bleeding in the gut and these may increase the presence of Hb in faeces. NSAIDs, such as aspirin and cyclo-oxygenase-2 enzyme (COX-2) inhibitors, may also, through irritation or changes to the intestinal microenvironment, increase the amount of Hb in the gut which may cause false positive test results. In the Scottish arm of the UK CRC screening pilot

participants taking regular anticoagulant therapy at the time of testing had an increased likelihood of a false positive FOBT result of 10%. This suggests that this group have an increased exposure to colonoscopy, a procedure with associated risks, with a negative outcome and no significant clinical findings. It was recommended by this group that, where possible medication was stopped three days before starting the gFOBT (95).

However, evidence of false positives due to taking medication was varied in a later systematic review conducted by Konrad et al., who observed that the standard of these studies did not allow for the evaluation of sensitivity and specificity of gFOBT for those taking anticoagulants and NSAIDs (96). A US study of 193 veterans reported no association between regular aspirin or NSAIDs use and false positive gFOBT results (97) Therefore, it appears that, there is no good evidence that long-term use of aspirin and NSAIDs would be linked to false positive gFOBT results. The pragmatic approach would be to not ask participants to stop taking medication which could potentially lessen uptake (uptake being one of the key factors in the effectiveness of a screening programme (94)) and this approach has been adopted in the SBoSP.

An early study in England reported that prescription of more than 5 years of aspirin or NSAIDs was strongly associated with reduced incidence of advanced adenomas (98). This has been replicated in a 20 year follow up of five large trials of aspirin use. Rothwell et al. reported, in a meta-analysis of these studies, a reduction in the incidence and mortality of CRC where doses of greater than 75 mg aspirin per day were being taken and it is likely that the causes of the reductions is the protective effect of long-term aspirin use rather than falsely elevated gFOBT results (99).

Rehydration affects outcome

If samples are rehydrated before testing, a higher number of positive test results without the commensurate increase in disease detection decreases specificity for CRC (28). Although, in the Minnesota RCT, rehydration of the gFOBT was associated with a reduction in long term CRC mortality (34), to avoid large numbers of referrals for colonoscopy, it has been recommended that a sample should be dried at least 48 hours on the gFOBT card before the developer is applied – this also minimises interference from plant peroxidases (90).

A consideration not highlighted by Beg (87) (Table 5) is improper storage of test cards prior to development. Improper storage conditions including exposure to excessive heat and light which renders alpha guaiaconic acid unstable. This increases the likelihood of the number of positive test results seen is increased with no increase in disease detection (82).

A false negative test result is a missed opportunity to visualise the bowel and detect underlying pathology. The most common factors leading to false negatives are:

High levels of vitamin C reduce positivity

Vitamin C (ascorbic acid) reduces the pseudo-peroxidase activity of haem and excessive amounts in the diet can cause false negative test results since the antioxidant action inhibits the production of the blue colour that indicates a positive result (87). With a number of studies reporting the inhibitory effect of doses of vitamin C in the range 1000 – 1500 mg per day,

it may be recommended that those wishing to participate in screening using a gFBOT refrain from taking vitamin C before collecting samples (93).

Improper sampling or development of gFOBT

In many screening programmes, the participant is responsible for taking the samples and returning the sample to the testing laboratory (56). The participant is required to smear two samples from three different faeces on the test card and date all of these correctly. Where the test card is not dated a decision cannot be made about the age of the samples and they must be as out-with the 21 day period of sample integrity. Analysis of gFOBT requires considerable training and expertise. It is subject to variations in the interpretation of subtle colour changes. Difficulty in interpretation of colour change, inexperience at point of care, incorrect sampling technique and insufficient time left for the sample to dry before development, can all lead to incorrect result interpretation (100).

Lesion does not bleed at time of faecal collection

The usual gFOBT test format requires the participant to apply faeces from one bowel motion to two sample windows on the card and do this for three consecutive bowel motions. This multiple sampling increases the probability that lesions that bleed intermittently or at low levels are detected. However, no screening method that detects Hb in the faeces as a marker of neoplasia will detect CRC that do not bleed. This group of CRC are not amenable to screening with faecal tests for Hb of any kind. There is a potential role for methods other than FOBT in screening for CRC as discussed in Chapter 1 (101).

Hb in faeces may be below the analytical detection limit of the test method

The gFOBT method does not give a positive result in all cases where there is Hb in faeces. The performance characteristics of a gFOBT are determined by the manufacturer. In the case of the hema-screen gFOBT a concentration of 600 µg Hb/g faeces is the analytical detection limit reported by the manufacturer. This means that some people who perform the test will receive a negative test result even when they do have Hb in their faeces that is related to CRC.

When advising participants about the sample collection procedure, it must be mentioned that the faecal samples are small enough (about the size of a pea) to dry on the gFOBT card prior to analysis; this acts as a bacteriostatic measure in the faeces. If this does not happen, false negative test results may be reported where the amount of Hb in the faeces has been near the analytical detection limit of the method and the action of faecal bacteria degrades haem to a concentration below the analytical detection limit of the gFOBT (87). That false negative test results occur with the gFOBT is evidenced by the high number of IC: 30 - 60 % of all CRC diagnosed in the screened population, in the Scottish Pilot rounds, were present in those who had had a negative gFOBT. These figures increased as more incidence rounds were completed (33).

Ultimately, the effectiveness a screening programme using gFOBT as the first-line test is in part determined by the number of participants who take part in screening and the number referred for colonoscopy; this in turn is dependent on the performance characteristics of gFOBT and minimisation of sources variation in the pre-analytical, analytical and post analytical phases.

2.4 Benefits and Disadvantages of Qualitative FIT

Qualitative FIT utilise the principle of the antibody-antigen reaction to directly detect human globin in faeces, thereby avoiding dietary sources of interference. This leads to a removal of one of the sources of false positives test results that compromised the effectiveness of the gFOBT. In addition FIT is more specific for lower GI bleeding than gFOBT as globin is more amenable to digestion by bacterial enzymes than haem. Globin is quickly degraded by bacteria in the intestine and so FIT are more sensitive for bleeding from the large bowel, rather than further up the gut. The location of bleeding may impact upon the concentration of Hb as bacteria degrade the intact Hb and a longer transit time potentially increases Hb degradation. This issue also appears to be sex related, with women often having a lower f-Hb than their male counterparts (14). Without a bacteriostatic preservative in the FIT sample collection device, microflora from the faeces may continue to degrade the globin, in some cases so much so that globin is not detectable (82). This issue is discussed further in Chapter 3.

Qualitative FIT are easy to use and the test can be performed by healthcare practitioners in the context of point of care testing (POCT). This type of screening model could provide fast turnaround of results and potentially avoid the additional costs of centrally organised screening. Qualitative FIT have overall improved performance characteristics than gFOBT, including better detection of adenomas, but like gFOBT offer only a positive or negative outcome. The analytical detection limits of the most commonly available qualitative FIT are more sensitive for Hb than gFOBT. This would result in more referrals for colonoscopy but only a slight increase in the number of CRC detected, thus the sensitivity is increased but with a loss of specificity.

Qualitative FIT are produced by many different manufacturers who, due to a lack of consensus, set different analytical detection limits. A recent study compared six qualitative FIT performed on patients who all had undergone colonoscopy (102). Here, the FIT performed better than a single gFOBT in all categories assessed: location, number of adenomas and size. However, a great deal of variation was found in the diagnostic performance of the different qualitative FIT and this suggests caution must be used when selecting a qualitative FIT method since different analytical detection limits impact on clinical outcomes when used in a screening context.

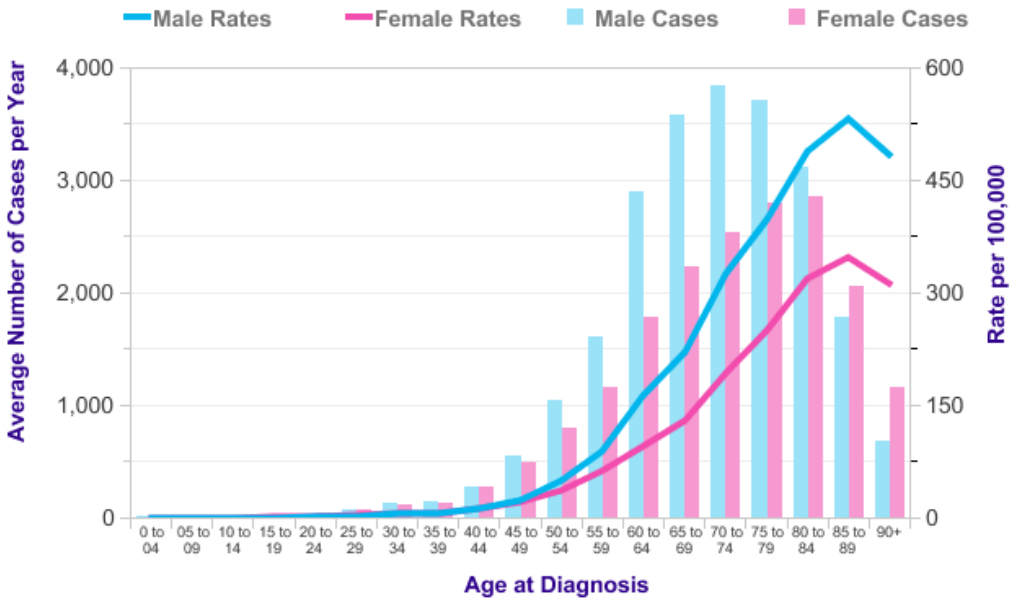
Further evaluation of the six different qualitative FIT methods gave positivity ranging from 6.4 – 46.8%, sensitivity 29.8 – 73.4% and specificity 58.8 – 96.7% (103). It can be seen from these results that different qualitative FIT methods do not give the same positivity, sensitivity or specificity when assessed against each other. Qualitative FIT give fewer false positives than gFOBT but they also have the potential to give different performance characteristics within the same test modality when applied to a screening setting and these are not within the control of the programme organisers.

The issue of apparent different performance in different manufacturers' products can be addressed by adopting standardised units relating to the mass of Hb in the mass of faeces. This requires a conversion factor from the concentration of Hb in the volume of buffer (ng Hb/mL buffer), for each manufacturer to micrograms of Hb per gram of faeces (µg Hb/g faeces) (104). Adoption of these units has been recommended by the Expert Working Group on FIT for Screening of the Colorectal Cancer Screening Committee of the World Endoscopy Organization (105) and would give programme organisers a better indication of performance across analytical methods (106).

2.5 Positivity when using gFOBT and Qualitative FIT in the Scottish Bowel Screening Programme (SBoSP)

Within the context of the UK, Scotland has the highest crude rate (number of new cases over a year per 100,000 of the population at risk) of CRC incidence - exceeding the four nations' average, which is shown in Figure 8 (2).

Figure 8 Incidence of CRC in UK



Cancer Research UK, Bowel cancer incidence statistics (2)

The Scottish Colorectal Cancer Screening pilot of screening was based on the Nottingham RCT of the 1990s. The pilot rounds of CRC screening invited men and women in three NHS Boards in Scotland from the age of 60 – 69 years to test using a gFOBT only algorithm (96). The results of the RCT (25) were mirrored in Scotland (46) and England (107). In addition, in 2013, a reduction in disease specific mortality of 10% (27% when corrected for participation) was reported from Scottish data using a matched cohort study comparing screened and unscreened areas during the pilot rounds of screening (5). The gFOBT used in Scotland is seen in Figure 9.

As a result of the pilot studies, programmes of CRC screening have been rolled out across the UK. The SBoSP currently invites all men and women between the age of 50 and 74 years (with the option to opt-in over the age of 75 years) to participate in screening using a two-tier reflex algorithm (66). Each eligible person is offered a gFOBT at invitation.

Figure 9 Scottish Bowel Screening Programme (SBoSP) gFOBT

Image courtesy of the Scottish Bowel Screening Laboratory (2014)

Any participant testing positive in 5 or 6 windows is immediately referred for colonoscopy in their local NHS Board (0.4%) and their GP notified. Those testing positive in 1 – 4 windows, i.e. any weak positive results (8.0%), are offered a qualitative FIT to complete. In this group approximately 20% have a positive test outcome. These participants are also referred for colonoscopy. The qualitative FIT used in Scotland is seen in Figure 10.

This algorithm has been in place since June 2007. Over that timeframe, the positivity and hence the referral rate, has been approximately 2.4% (set against a projected positivity of 2.1%). Uptake within the SBoSP has consistently been 53 – 54%. Participants are eligible for invitation once every two years irrespective of previous outcomes. Participants are

excluded when their GP confirms they have no colon or if they wish to opt out of screening.

Figure 10 SBoSP qualitative FIT

Image courtesy of the Scottish Bowel Screening Laboratory (2014)

Clinical outcomes are captured nationally in the Key Performance Indicators (KPI) published by Information Services Division Scotland (ISD). The most recent publication (108) shows that more women than men participate in the SBoSP and that uptake decreases with increasing deprivation. The overall referral for colonoscopy is 2.4% and, of those referred, the crude CRC detection rate is 0.20% for men and 0.10% for women. The adenoma detection rate is 1.08% and 0.42% for men and women respectively. The overall positive predictive value (PPV) for polyp CRC or invasive CRC in those who have had a positive screening result and proceeded to have a colonoscopy is 6.4%.

2.6 Defining a Quantitative FIT Algorithm

The Expert Working Group (EWG) on FIT for Screening, on behalf of the Colorectal Cancer Screening Committee (CRCSC) of the World Endoscopy Organisation (WEO), determined that there were 13 automated quantitative

FIT analytical platforms that could provide an estimate of quantitative f-Hb (84). It is of note that in this group there was one manufacturer that supplied four different analytical platforms to suit a range of healthcare settings, from point-of-care to large centrally organised laboratories.

The majority of automated quantitative FIT available utilise a small specimen collection device which collects faeces in a serrated or spooned probe attached to the cap of the device. These collect between 2 and 10 mg of faeces into the body of the device which contains between 1 and 2 ml of buffer and matched bench-top analyser to provide the testing system. Sample collection devices and analysers are not generally interchangeable. Such systems are currently being trialled or are in use in many European countries. Despite it appearing that FIT could be used as a direct replacement of the gFOBT as a test for FOB in faeces, there are many variables that need to be considered before the introduction of FIT in a screening programme. These include, but are not limited to – number of samples used, screening interval and f-Hb cut-off: each is examined in further detail below.

A FIT approach requiring a single sample might well be associated with a higher uptake than the two-tier approach currently used in Scotland which requires three samples for the gFOBT and two for the qualitative FIT (109). Concerns regarding the reduction of the number of samples used in screening are that, of the CRC that bleed, some may bleed intermittently or the blood may not be distributed evenly across the faeces passed. This issue has been addressed in a number of studies which are described below.

In The Netherlands, a study of 1,096 participants who collected two faeces samples and had colonoscopy based on three possible outcomes (positive/negative, negative/positive) (positive/positive) (mean of both FIT) at increasing f-Hb cut-off. These concentrations were compared to disease

severity. This study strongly suggested that double FIT sampling does not provide a superior combination of sensitivity and specificity compared to single FIT sampling. On the other hand, it indicated that an increase in sensitivity could be achieved by lowering the f-Hb cut-off for reporting of a positive test result rather than increasing the number of samples required (110).

In another study, 1,682 patients with either normal or slightly elevated risk of CRC, attending one of three clinics for colonoscopy in Israel, provided further information regarding cut-off and number of samples to use. Each participant completed samples from three different faeces and returned them to a central laboratory for analysis. It was shown that, using the lowest f-Hb cut-off, that using 'any sample as positive' as the criteria gave 100% sensitivity for CRC. The sensitivity with one FIT at the same f-Hb cut-off was 75%. Further analysis showed that the gain in sensitivity to CRC and advanced adenomas comes at the cost of specificity; the low f-Hb cut-off meant that the numbers sent for colonoscopy increased (111).

Another important study calculated sensitivity and specificity of the OC-Sensor Diana analyser (Eiken Chemical Co., Tokyo, Japan) in a mixed cohort (screening, surveillance and symptomatic) at six increasing cut-off from 50 - 200 ng Hb/ml buffer (112). At the lowest cut-off (50 ng Hb/ml) the test detected 84.2% of early CRC and at the highest cut-off (200 ng Hb/ml) 78.9%, but there was a larger decrease in detection of all screen relevant neoplasia, from 47.1% to 37.2%. The number of positive test results dropped from 16.5% at 50 ng Hb/ml buffer to 10.2% at 200 ng Hb/ml buffer. The authors propose that cut-off can be adapted to suit resources, suggesting that, in the prevalent round, a high cut-off could be used as most prevalent CRC will be detected and that as more incident and less advanced CRC form in the population, the cut-off can be reduced making the test more sensitive for early disease than using the cut-off of the prevalent round.

There has been concern that Hb is not stable enough in native faeces to be used as part of a screening programme in which sampling devices are posted to a central laboratory for processing (113). Drying faeces, as described in the gFOBT method, reduces this effect and card based FIT sample collection systems are available (90). Many FIT technologies use a liquid-based sample collection system and so the buffer must contain preservatives and stabilisers that will not interfere with the test.

Recent studies have investigated the stability of Hb in buffer over time and at different temperatures. Dutch studies describe the average return time for one FIT in their screening trials as three days. They tested 71 positive FIT every three to four days for a further three weeks and plotted the Hb concentration. They conclude that none of the positive test results suffered from Hb degradation to such an extent that the test results changed from positive to negative within 10 days and that delays in returning samples to the laboratory of up to one week would not necessitate repeat sampling. Over this time, however, Hb concentration may be reduced in the FIT sample collection device which may have a clinical impact (114). This is in line with the Dutch findings (115) that delays in testing may cause false negative test results.

A study in the Florence district of Italy (116) determined f-Hb from the screening programme and aggregated them into four seasons – spring, which had a mean temperature 27.6^{0}C (95% CI 26.2 to 29.1); summer 25.2^{0}C (95%CI 23.1 to 27.3); autumn 29.2^{0}C (95% CI 27.7 to 30.6); winter 29.5^{0}C (95% CI 27.9 to 31.1). These data were analysed to determine the impact of temperature on FIT positivity. They found that there was a decrease in positivity during the warmer months and proposed that this might affect the use of FIT, particularly in countries with high ambient temperatures. These issues must be addressed by manufacturers and

newer generations of FIT buffers should be developed to render Hb more stable and more suitable for use in population screening.

Many studies investigated development of an evidence base for the use FIT algorithms have used one, two or three sampling devices at varying f-Hb cut-off and with varying intervals. Currently, there are no firm recommendations within the literature regarding which is the best strategy to adopt. However, it is clear from these studies that the automated, quantitative FIT algorithm used in a national screening programme would be adaptable according to the constraints of available resources.

2.7 Comparisons of gFOBT and FIT

The diagnostic accuracy of one gFOBT, the Stool Occult Blood Test Kit (Cenogenics Corporation, Morganville, New Jersey) and two FIT analytical systems, the OC-Sensor μ (Eiken Chemical Co., Ltd, Tokyo, Japan) and FOB Gold (Sentinel Diagnostics, Milan, Italy) have also been assessed (117). In this study, three faecal samples were initially assessed by gFOBT and FIT to determine that they tested as negative for f-Hb. One sample was retained as a negative result and two lysates prepared from heparinised human blood samples were used to give f-Hb of 2.5 mg Hb/g faeces (weak positive test result) and 4.5 mg Hb/g faeces (strong positive test result). They asked nine experienced ward staff to prepare and test one sample from each of the faecal samples on gFOBT cards and tested them on two quantitative FIT systems with the cut-off at 100 ng Hb/ml buffer. The ward staff gave a correct reading for the negative sample in eight out of nine tests; they reported a correct reading in five of nine samples in the weak positive sample and seven out of nine in the strong positive sample. The FIT determined all of the negative faeces as negative and all of the spiked samples as positive. Overall, the analytical sensitivity and specificity of the

gFOBT were 67% and 89% respectively, the quantitative FIT outcomes were both 100%.

Evidence relating to the clinical advantages of using FIT compared with gFOBT in the detection of CRC derive from a number of studies, the most compelling of which comes from The Netherlands where gFOBT and an automated, qualitative FIT (OC-Sensor μ) were compared in a randomised population-based trial (118). The FIT yielded a significantly higher PPV for both CRC and advanced neoplasia, a significantly lower false positive test percentage and a higher uptake.

These results are not directly applicable to the SBoSP since, using the manufacturer's recommended f-Hb cut-off the positivity for the FIT was 5.5%, too high for the existing colonoscopy capacity in Scotland. However, this study did demonstrate unequivocally that gFOBT significantly underestimates the prevalence of CRC and adenoma in a screening population when compared with FIT under the conditions adopted in the study. Further work from France, in which participants carried out both a gFOBT and a FIT simultaneously, produced a similar endorsement of FIT as the better test (119).

2.8 Benefits of Automated Quantitative FIT

Earlier in this Chapter, the disadvantages of gFOBT when screening for CRC were examined and how qualitative FIT addresses some of these was discussed. This section outlines how many of the remaining disadvantages of qualitative FIT are addressed by the quantitative FIT option (100).

In 2012, the FIT Guidelines Expert Panel in Ontario, Canada (120), published a comprehensive review of published literature detailing most relevant gFOBT/FIT comparisons, i.e. those systems that are approved for processing in laboratories in Canada. It concluded that the ideal FIT would have the characteristics described in Table 7.

Table 6 Characteristics of the ideal FIT

1	Provide a numerical result (so the f-Hb cut-off can be set by end user)
2	Be readily automated in the laboratory
3	Require one faecal sample
4	Have specimen stability across a wide variation in temperatures
5	Have sample stability of at least 7 days from date of collection to date of analysis

FIT Guidelines Expert Panel Rabeneck et al (120)

A number of evaluations of automated quantitative FIT platforms have been undertaken. Vilkin et al. undertook an evaluation offering a three FIT sample regime, tested on the OC-Sensor analyser, to a symptomatic cohort of patients presenting in primary care. It evaluated test reproducibility, Hb degradation, intra-patient variability of f-Hb, sensitivity and specificity for significant neoplasia (121). This study found that five prepared faecal sample collection devices tested five times did not produce significantly different results. When 30 samples were stored at 4^0C, 20^0C and 28 – 30^0C for three weeks, those stored at 4^0C did not have f-Hb that was significantly lower than their original concentration, but those stored at 28 – 30^0C did demonstrate a significant drop in f-Hb. In each category of disease, f-Hb increased with disease severity. The study reported an additive clinical value of repeated FIT with a sensitivity of 76.5% and specificity of 95.3% at 100 ng Hb/ml buffer. In this study, even though the FIT was more sensitive than the gFOBT for significant neoplasia, because of the f-Hb cut-off and the high number of referrals for colonoscopy, they did not find it more sensitive than a cheaper gFOBT.

Generally one sample is required

It has historically been that gFOBT used in screening programmes have been performed on three consecutive samples. Two studies have investigated the effectiveness of one v two FIT sampling regimes. In the Dutch study all invitees for colonoscopy, were asked to complete one or two FIT samples (Eiken OC Sensor). Results were analysed in three ways; one of two FIT positive, two of two FIT positive and where the mean of two FITs exceeds the positive cut-off. This was performed across a range of cut-off values 50, 75, 100, 150 and 200 ng Hb/ml buffer. The second study was conducted on a screening population across four areas of Italy. Participants with at least one result >79 ng Hb/ml buffer were referred for colonoscopy, those with a result < 80 ng Hb/ml buffer received a negative result letter. In both studies the authors found that using two FIT over one FIT at different cut-offs did not improve the sensitivity and sensitivity of the test and that if sensitivity, for CRC or advanced adenomas was the main concern this could be improved by using one FIT with a low cut-off (110) (122). There is considerable evidence that this approach with FIT is associated with higher participation than the current gFOBT approach (118) (123).

Screening interval can be extended

Current evidence suggests a 10 year development process from adenoma to carcinoma. This is supported by histopathological data from adenomas left *in situ* (11) and more recent genetic and cytogenetic evidence (8). Based on this, biennial screening should detect early CRC and adenomas before symptoms appear. The RCT described in the previous Chapter where the decision was taken to use two year (and in the case of the Minnesota RCT one year) screening intervals demonstrated a significant reduction in CRC mortality (25). However, the introduction of quantitative FIT set at a lower f-

Hb cut-off could provide an opportunity for extending the interval between screening rounds without reducing the PPV for CRC.

A Dutch study, inviting over 10,000 screening naïve people, to undertake one FIT test at one, two or three year intervals over two screening rounds, as compared to a reference group who were offered once only FIT, reported interesting outcomes. It was determined that the uptake of a second invitation was higher in the two and three year interval groups than those invited annually. In addition, the number of individuals with advanced neoplasia was not influenced detrimentally by the length of time between FIT. The implication is that screening intervals up to three years could be used to tailor referral rate for colonoscopy without any detriment to neoplasia detection and that uptake may improve with longer screening intervals (124).

<u>f-Hb cut-off can be set by the user, dependent on resources</u>

The use of automated quantitative FIT is becoming widespread and seen by many as the current method of choice (27) (31). However, one major drawback is that manufacturers often recommend f-Hb cut-off that causes referral to colonoscopy to be 10 – 16%. This would be too high for UK countries to cope with at the present time. The approach of using one f-Hb cut-off uses the quantitative FIT simply as a qualitative investigation, i.e., dividing the participants into two classes, those who do not and those do warrant further investigation, usually colonoscopy. As numbers referred increase, sensitivity for neoplasia increases but specificity decreases as more people have to undertake colonoscopy to find neoplasia (125). Where colonoscopy capacity is a major consideration, it would be a benefit for the programme to be able to adjust the f-Hb cut-off in line with local colonoscopy resources, decreasing the f-Hb cut-off in successive rounds of screening.

Aspirin use can be beneficial

There is evidence that, when undertaking screening with the newer FIT modality, the use of low-dose aspirin is associated with a markedly higher sensitivity for detecting significant neoplasia (126) (127). This was seen in two different automated quantitative FIT systems evaluated and was not seen with gFOBT. It may be that, in people with smaller lesions, taking aspirin causes increases in gastrointestinal blood loss that would be undetectable using gFOBT, have detectable f-Hb using FIT. A further study found that aspirin use was not a risk factor for causing false positive test results (128).

FIT cost more than gFOBT at the point of invitation to screen, but, there are very few direct evaluations of the cost benefits of gFOBT and FIT. The Centre for Reviews and Dissemination reported on seven economic evaluations: the extra cost of FIT is likely to be cost-neutral or cost-saving to the potential for an increase in the detection of significant lesions and reduction in chemotherapy and treatment costs further on in the care pathway (129) (130).

With automated quantitative FIT there now exists the technology to determine f-Hb. There is growing evidence to support the utility of this approach in determining not just the presence of CRC, but adenoma detection within screening programmes.

2.9 Conclusion and Aims of This Work

Automated quantitative FIT technology has now become widely accessible? FIT have many advantages and some disadvantages when compared to gFOBT as an examination to be used in CRC screening programmes.

There are a number of areas that would benefit from further study and these form the main questions put forward in this thesis:

- can FIT be successfully introduced into the current SBoSP,
- does FIT outcome have a relationship with demographic variables, and
- can FIT be used to triage symptomatic patients?

Since FIT are highly specific for Hb and its early degradation products, they offer the possibility of reducing false positive test rates in bowel screening programmes leading to fewer unnecessary colonoscopies. This provides opportunities for optimising demand on colonoscopy services and delivery improved population outcomes. The availability of quantitative data allows the programme organiser, not the manufacturer, to set the f-Hb that divides positive from negative test results [cut-off] and thereby enables adjustment of the positivity to suit endoscopy capacity.

Colonoscopy is a scarce resource. Any test that can prioritise those who need urgent referral for colonoscopy would be of great benefit to the patient and the health care system. The FIT could be performed on a symptomatic patient presenting in primary care and the f-Hb be used to decide the appropriate care pathway. The algorithm developed for the triage of symptomatic patients would differ from that used in a screening setting since sensitivity would ideally have to approach 100% if used as a rule-out diagnostic test. Evidence must be gathered in order to demonstrate that very few patients with CRC would be missed by this strategy.

This Chapter has set out the growing evidence that detection of CRC would be best performed by automated quantitative FIT using immunoturbidimetric analysis. The next Chapter details the verification of the FIT analyser used for the work described in this thesis.

CHAPTER 3: Evaluation of a Quantitative FIT Analytical System

3.1 Introduction

Currently there are several automated analytical systems that can provide f-Hb measurement. Automation enables movement from manual processes and subjective result interpretation which cast limitations on both gFOBT and qualitative FIT as described in Chapter 2. This Chapter outlines the validation and verification of the quantitative FIT analyser used in this work and the methods used to confirm or refute the manufacturers stated performance characteristics.

Manufacturers provide information regarding the specification of their product. Details will vary but should include: linear range, number and concentration of control materials, reported range of measurement, analytical detection limit, imprecision and Hb stability. A number of these characteristics are explored further in this Chapter.

The current interest in automated FIT analysers prompted an evaluation of the four most widely used analytical systems considered suitable for use in the NHS Bowel Cancer Screening Programme, and by implication SBoSP, by the Guildford Medical Device Evaluation Centre (GMEC) (131). Detailed analysis of the technical performance and suitable specimen collection systems contained in the report helped inform the choice of analyser.

3.2 Characteristics of the Analytical System Adopted

The OC-Sensor Diana analyser is a bench-top (630 mm x 560 mm x 560 mm and 60 kg), automated quantitative analytical system that replaces the smaller OC-Sensor µ. The O-C Sensor Diana was reported to have the capacity to test 280 samples an hour and be able to continuously accept

new samples. It is able to hold two bottles, each containing 200 ‘shots’, of latex; these can be replenished during analysis.

The OC-Sensor Diana was the only analyser to be CE-marked for use in the UK at the time of the work described here: use of such certified methods is essential for the maintenance of external accreditation (the SBoSL currently holds unconditional accreditation with Clinical Pathology Accreditation UK Ltd (CPA)) (132). It is one of the Quality Improvement Scotland standards that the SBoSL hold accreditation to ISO 15189 standards (133).

The OC-Sensor Diana is the analyser from which much of the evidence on the merits of FIT have been determined in Europe.

As described already, during periods of this work two OC-Sensor Diana analysers were used, this was to provide a backup should one fail during the evaluation of FIT in the SBoSP. The positioning of the analysers is shown in Figure 11. A representative from MAST Group Ltd (MAST House, Bootle, Merseyside, UK) was assigned to support the laboratory and maintain the analysers.

Figure 11 Set up of two OC-Sensor analysers

Image courtesy of the Scottish Bowel Screening Laboratory (2014)

At the time of this work there was no World Health Organization (WHO) reference standard suitable for demonstrating metrological traceability for immunochemical tests with low analytical detection limit for f-Hb. OC-Sensor standards and controls are calibrated using an internal standard which is calibrated to hemoglobincyanide 98/708 National Institute for Biological Standards and Control (NIBSC) (134). This was still the case after the work was completed.

3.2.1 Verification of the Method

Introduction of any new analytical system within the SBoSL requires demonstration that the method is suitable for its intended use (135). This includes evaluation of the pre-analytical process, analysis of samples and post-analytical steps. In addition there must be evaluation of the analyser *in situ* forming a comparison with the manufacturer's stated analytical performance characteristics where available. All these facets were investigated during the validation and verification of the analyser. This followed the Department of Biochemical Medicine, Ninewells Hospital and Medical School, Standard Operating Procedure (SOP) - Evaluation of a new method and are included in the SOP SBSL 31 How to operate the OC-Sensor analyser (Appendix 1) written before the work commenced.

The study undertaken within the symptomatic setting commenced February 2010 and ended March 2012. The period of study involving participants in the SBoSP started July 2010 and ended April 2011. It was only during the period when testing samples for the SBoSP that two analysers were used. During this part of the work, each analyser was set up on alternate days: if the chosen analyser was out-with the control limits set for acceptance of the run, the alternate analyser was set up and run.

3.2.2 Analysis of Intra-run and Inter-run Imprecision

The imprecision of a method (dispersion) is usually measured using standard deviation (SD), whilst the magnitude of dispersion is interpreted by using the coefficient of variation (CV). Intra-run variation is determined by testing the same sample/control material repeatedly on the same day. Inter-run imprecision is determined by testing freshly prepared samples/control material over a number of days. The ideal minimum number of replicates for the determination of standard deviation is 20 and it is also useful to use a variety of sample types. In this instance there was no opportunity to use patient samples as the analyser was not in use anywhere else in Scotland and so high and low control materials were used alongside 'spiked' faecal samples.

The high and low control material was supplied by the manufacturer as lyophilised powder in vials and required the addition of 1.0 ml of purified water and to be allowed to stand for 15 minutes before use. The manufacturers high and low control ranges were LOT specific. The other type of material used in the assessment of imprecision was native faeces 'spiked' with human whole blood lysate which had been diluted to a concentration that was within the analytical range of the analyser. The manufacturer provided all the other components required for testing, including latex, calibrator, buffer, diluent and bleach.

Initially only one analyser was *in-situ*. High and low control material and four 'spiked' samples were used to determine intra-run variation. To prepare these, fresh specimens of faeces from four apparently healthy volunteers were tested for haemoglobin with hema-screen SPECIFIC. They were all negative. After thorough mixing, one portion of each was saved; the rest were supplemented with Hb lysate from a redundant pack of SNBTS blood. This was prepared by freezing 1 ml aliquots and thawing for use. Four different concentrations of Hb lysate were used, giving five sets of results. The range of concentrations used reflects clinically relevant concentrations

of Hb. Due to the heterogeneous nature of faecal material, distribution of the lysate was difficult, accurate sampling into the collection tube also proved to be difficult as the serrated end of the sampling stick is small. Given these factors, the initial Hb concentrations show variation within the four spiked groups. Due to the difficulty of preparing faeces and Hb lysate at this early stage the spiked samples all have low levels of Hb present.

However, given the constraints, the analyser was shown to have similar imprecision to that stated by the manufacturer within the linear range 40 – 1000 ng Hb/ml buffer. The raw data is shown in Table 7. In the case of Sample 1, it is usual for measurements near the analytical detection limit of a method to have greater dispersion and higher coefficient of variation. During the course of the evaluation samples with concentrations greater than the upper limit were not diluted and re-assayed.

Table 7 OC-Sensor spiked samples run on one analyser over one day

4 spiked (Hb lysate) faecal samples

High control material (mean) 630ng Hb/ml buffer,

Low control material (mean) 155 ng Hb/ml buffer

REPEAT	SAMPLE 1	SAMPLE 2	SAMPLE 3	SAMPLE 4	LOW	HIGH
1	16	32	64	74	162	669
2	21	37	70	76	165	685
3	29	44	77	78	164	688
4	15	43	74	84	171	701
5	26	44	69	81	176	758
6	30	49	77	88	176	701
7	33	53	76	81	173	713
8	25	54	73	86	175	713
9	24	48	73	94	173	716
10	35	59	71	89	167	698
11	30	55	72	88	169	706
12					174	702
13					175	724
14					182	728
15					180	724
n	11	11	11	11	15	15

Low	15	32	64	74	162	669
High	35	59	77	94	182	728
MEAN	25.8	47.1	72.4	83.5	172.1	708.4
SD	6.5	8.1	3.9	6.1	5.8	21.1
CV (%)	25.2	17.2	5.4	7.3	3.4	3.0

When the second analyser was validated the imprecision studies were revisited and further high and low control material were run fifteen times on one day on both analysers named OC-Sensor 1 and OC-Sensor 2 (new analyser). The raw data concerning intra-run imprecision on two analysers are contained in Table 8.

Table 8 OC-Sensor intra-run imprecision

High control material (range) 144 - 168 ng Hb/ml buffer,

Low control material (range) 573 - 681 ng Hb/ml buffer

	OC 1		OC 2	
REPEAT	LOW	HIGH	LOW	HIGH
1	147	628	143	594
2	149	637	145	599
3	148	633	144	603
4	149	625	142	605
5	151	629	145	608
6	150	634	142	572
7	142	595	142	596
8	148	612	142	596
9	149	610	142	597
10	148	621	146	603
11	150	633	149	605
12	152	641	149	607
13	151	627	151	588
14	154	700	153	610
15	155	647	153	625
n	15	15	15	15
Low	142	595	142	572
High	155	700	153	625
MEAN	151	631	146	601
SD	3.1	23	4.1	12
CV (%)	2.0	3.6	2.8	1.9

Limited variability in reported results is expected in these intra-run data and this is reflected in the low coefficient of variation (CV): no further statistical tests were performed on these data.

Assessment of inter-run imprecision required that the shut-down, restart and daily maintenance was performed according the SOP between each set of controls being tested. These data are shown in Table 9.

Table 9 OC-Sensor inter-run imprecision

High control material (range) 144 - 168 ng Hb/ml buffer

Low control material (range) 573 - 681 ng Hb/ml buffer

OC 1

Day	LOW	HIGH
1	140	623
2	145	621
3	168	700
4	145	596
5	142	598
6	150	622
7	148	618
8	150	620
9	151	625
10	148	616
n	10	10
MEAN	148.7	623.9
SD	7.7	28.6
CV (%)	5.2	4.6

OC 2

Day	LOW	HIGH
1	158	671
2	163	616
3	164	615
4	154	597
5	154	605
6	149	603
7	149	608
8	150	610
9	146	600
10	152	644
n	10	10
MEAN	153.9	616.9
SD	6.1	23.1
CV (%)	3.9	3.7

There are two statistical tests that could be used to assess the data from two analysers. The t-test was used in a non-parametric, comparison of two groups of continuous variables to assess the difference of the two means for the high and low control values across the two analysers. For the low control the P-Value equals 0.1100 which is not significant at $p < .05$ and the high control material, the P-Value is 0.5548, this is not significant at $p < .05$.

These results indicate that the high and low control values are not significantly different between the two analysers.

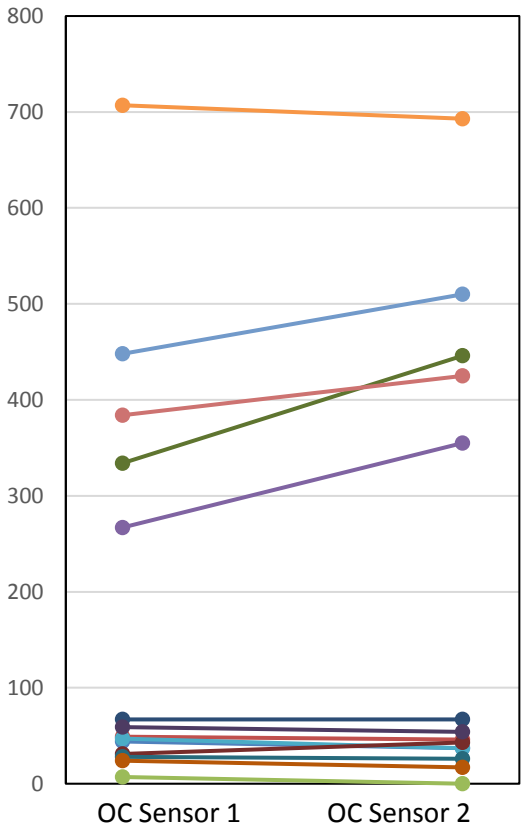
The F test allows comparison of the variance of two systems. Analysis of the low control material, gave F as 1.594: the critical value of F for $P < 0.05$ with 9 degrees of freedom [n-1] is 3.18 and for $P < 0.01$, the critical value is 5.35. So, F is less than the critical value, therefore, the imprecision of the two analysers at the low concentrations are not statistically different.

For the high control material, F was 1.778: the same critical values apply, so again there is no difference. Therefore, the data collected over 10 days show that statistically the analysers do not have different analytical imprecision.

Preparation of the samples for Hb degradation studies provided a small number of real faecal samples spiked with Hb lysate which could be analysed. These samples were used to set up a comparison across the two machines. These samples were run once on each analyser and the results are shown in Table 10 as raw data and a comparison table. This data shows greater parity in repeat tests with lower Hb concentrations than high faecal Hb concentrations.

Table 10 Real faecal samples measured on OC 1 and OC 2

SAMPLE	OC1	OC2
1	155	149
2	37	44
3	46	49
4	0	7
5	355	267
6	37	47
7	693	707
8	67	67
9	43	31
10	446	334
11	54	59
12	26	28
13	17	24
14	510	448
15	425	384
n	15	15
low	0	7
high	510	448



There is again an issue here regarding the heterogeneity of faecal samples and dispersion of faecal material and Hb. During the preparation of samples this is addressed by thorough mixing but may still be causing an artefact in the results.

3.2.3 Internal Quality Control Data

High and low control material was supplied by the manufacturer and run with participant samples at a rate of one high and one low control material for every 100 participant samples, or where there were fewer than 100 participant samples one set of high and low controls were analysed. This is similar to the internal quality control (IQC) process already used in the SBoSL with gFOBT and qualitative FIT. These data were checked to be

within limits and each run signed off as acceptable by Health and Care Professions Council (HCPC) registered Biomedical Scientist staff. Over the course of this work these data were collected and analysed. The imprecision, taking account of both analysers, during use in the evaluation period was less than 3.8%. These results reflect the manufacturer's claims and so were accepted for this method. These data are shown in Table 11.

Table 11 OC-Sensor 1 and 2 imprecision during evaluation period July 2010 – January 2011

			Low Control				High Control			
Date	LOT	OC	n	Mean*	SD	CV %	n	Mean*	SD	CV %
06/07	01001	1	38	154.7	8.35	5.41	38	637.1	31.81	4.99
30/08	01001	1	74	150.7	5.87	3.90	74	620.9	23.06	3.71
01/11	08001	1	10	147.9	2.47	1.67	10	642.3	19.32	3.01
13/11	08001	1	22	153.6	5.23	3.41	21	640.9	20.73	3.23
03/12	08001	1	23	146.0	5.01	3.43	23	598.8	22.29	3.72
24/12	09013	1	5	162.2	5.85	3.61	5	682.8	13.21	1.94
06/01	09013	1	7	155.7	3.15	2.02	8	676.0	24.17	3.58
12/01	09013	1	53	158.8	13.43	8.46	54	672.1	35.61	5.33
12/07	01001	2	61	153.0	15.21	9.94	61	622.3	27.49	4.42
31/08	01001	2	72	153.2	4.73	3.09	72	614.5	20.36	3.31
01/11	08001	2	9	151.2	3.53	2.33	9	610.8	24.49	4.01
09/11	08001	2	21	155.8	4.83	3.01	20	619.4	21.90	3.54
03/12	08001	2	26	152.0	5.36	3.53	26	636.7	37.33	5.86
24/12	09013	2	3	155.7	19.63	12.31	3	658.3	70.50	10.51
06/01	09013	2	7	160.7	3.95	2.46	7	702.6	11.90	1.56
12/01	09013	2	59	158.7	15.95	9.98	59	653.8	25.99	3.89

Date	LOT	OC	n	Mean*	SD	CV %	n	Mean*	SD	CV %
06/07	01001	1	106	152.4	6.5	4.3	106	627.7	25.9	4.1
01/11	08001	1	48	150.7	5.6	3.7	47	630.9	26.2	4.2
24/12	09013	1	52	160.5	5.1	3.2	62	672.5	24.4	3.6
12/07	01001	2	130	152.4	4.6	3.0	130	617.9	20.0	3.2
01/11	08001	2	54	153.6	4.8	3.2	54	625.5	27.5	4.4
24/12	09013	2	54	162.6	6.0	3.7	53	697.3	20.8	3.0
Assigned	01001			156	6			627	27	
Assigned	08001			154	5			626	25	
Assigned	09013			161	7			685	20	
Overall		1	206			3.9	215			4.0
Overall		2	238			3.2	237			3.5
Overall		1+2	444			3.5	452			3.8

*mean is reported as ng Hb/ml buffer **weighted for each LOT

The analysers were calibrated every month, or each time the latex LOT changed, whichever occurred first. This was followed by a period of accepting internal quality control values based on the manufacturers acceptance criteria. After 10 replicate analyses a local multirule acceptance criteria was calculated based on results being within two standard deviations of the mean (1_2s). This rule was used to determine acceptance or rejection of subsequent analytical runs.

3.2.4 External Quality Assessment Scheme Data

It is a requirement of the ISO 15189 standards that laboratories must participate in an External Quality Assessment Scheme (EQAS), where it is available, for each measurand that is examined in that laboratory. There is currently no EQAS scheme within the UK for FIT testing. However, the manufacturer of the OC-Sensor Diana analyser do organise a Japanese-based worldwide EQAS.

During the period of the work in which two analysers were in use there was one opportunity for the SBoSL to participate in this scheme. The manufacturer provided Sample A and Sample B, pre-prepared as liquid controls, ready to be placed in a control cup for analysis – this is similar to testing control material rather than patient samples.

Figure 12 Eiken external quality assessment report OC-Diana 1 (reproduced with kind permission from Mast Group)

IVD solutions through partnership

Mast Group

Mast Group Ltd.
Mast House
Derby Road
Bootle
Merseyside L20 1EA
United Kingdom

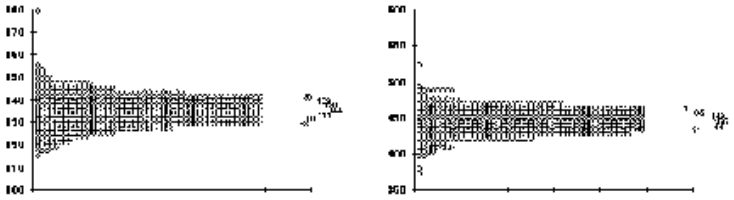
Tel. +44 (0)151 933 7277
Fax +44 (0)151 944 1332

www.mastgrp.com

Paula McDonald
Laboratory Team Leader
Scottish Bowel Screening Programme
Kings Cross Hospital
Clepington Road
DUNDEE DD3 8EA

2010 EQCS-OC

Thank you for taking part in this year's quality assessment programme.
Your results are detailed below.



Sample A		Sample B	
No of facility	805	No of facility	805
Mean value	135.5	Mean value	445.5
SD	5.7	SD	14.2
Max value	114	Max value	370
Min value	180	Min value	521

Analyser: **SN N000417 (OC Diana 1)**
Sample A: **131**
Sample B: **421**

Iain McElarney
New Product Sales Development Manager

Registered in England 632512
Mast House, Derby Road Bootle, Merseyside L20 1EA, United Kingdom

When testing Sample A, the mean attained by the 805 participating laboratories was 135.5 ng Hb/ml buffer. The result generated by OC-Sensor 1 was 131 ng Hb/ml buffer and by OC-Sensor 2 138 ng Hb/ml buffer. For Sample B, OC-Sensor 1 returned a result of 421 ng Hb/ml buffer and OC-Sensor 2 a result of 434 ng Hb/ml buffer, the mean for the group was 445.5

ng Hb/ml buffer. Analyser 1 gave a low result >1 SD away from the mean of the group and a high result < 1 SD from the mean. Analyser 2 gave results < 1SD away from the mean.

Figure 13 Eiken external quality assessment report OC-Diana 2 (reproduced with kind permission from Mast Group)

IVD solutions through partnership

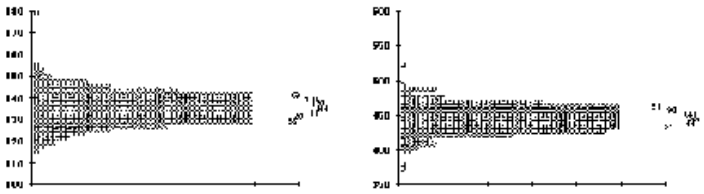
Mast Group

Paula McDonald
Laboratory Team Leader
Scottish Bowel Screening Programme
Kings Cross Hospital
Clepington Road
DUNDEE DD3 8EA

Mast Group Ltd.
Mast House
Derby Road
Bootle
Merseyside L20 1EA
United Kingdom
Tel. +44 (0)151 933 7277
Fax +44 (0)151 944 1332
www.mastgrp.com

2010 EQCS-OC

Thank you for taking part in this year's quality assessment programme.
Your results are detailed below.



Sample A		Sample B	
No of facility	805	No of facility	805
Mean value	135.5	Mean value	445.5
SD	5.7	SD	14.2
Max value	114	Max value	370
Min value	180	Min value	521

Analyser: **SN N000279 (OC Diana 2)**
Sample A: **138**
Sample B: **434**

Iain McElarney
New Product Sales Development Manager

Overall the results from the two analysers were low compared to the group. However, all returned results were within 2 SD of the mean. This confirms the results seen in the SBoSL evaluation of imprecision between the two

analysers and manufacturers claims. The Eiken reports are shown in Figures 12 and 13.

3.3 Hb stability

There is evidence that, when faeces are collected with no preservative present, degradation of Hb continues after collection (113). Manufacturers are unable to offer definitive guidelines regarding the stability of Hb in their sample collection devices due to a number of factors, namely each individual's faecal flora affects Hb degradation, increased transit time is likely to reduce the amount of detectable Hb and delays before testing may also contribute to a reduction in f-Hb. These issues cannot be replicated by the manufacturer. Manufacturers have suggested that users investigate local conditions.

The storage temperature and time taken to transport participant samples cannot be controlled. In order to investigate the effect of different temperatures and varying lengths of time before testing on f-Hb, 55 faecal samples were set up to mimic possible pre-analytic conditions.

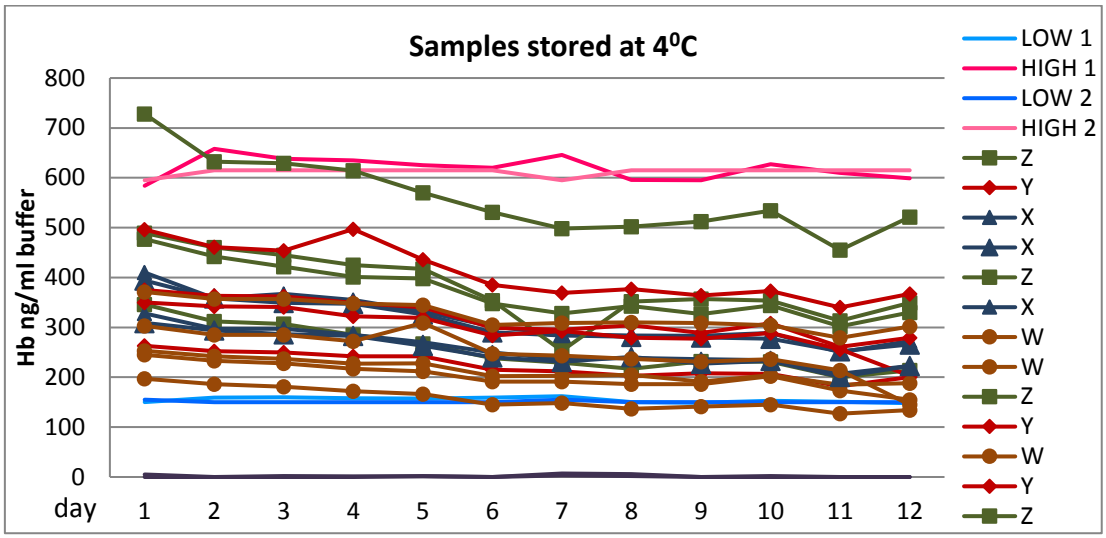
Fresh specimens of faeces from four apparently healthy volunteers were tested for the presence of Hb with hema-screen SPECIFIC (Immunostics, Ocean, New Jersey, USA). They were all confirmed to have no detectable f-Hb by analysis using the OC-Sensor Diana. After thorough mixing, one portion of each was saved. Hb lysate was prepared by freezing and thawing a specimen of venous blood from a redundant pack of blood supplied by Scottish National Blood Transfusion Service (SNBTS). Four different concentrations (using doubling dilutions) of Hb lysate were used with the portion that tested negative, giving five sets of f-Hb. The range of concentrations prepared reflected clinically relevant concentrations of Hb in faeces. Due to the heterogeneous nature of faecal material, incorporation

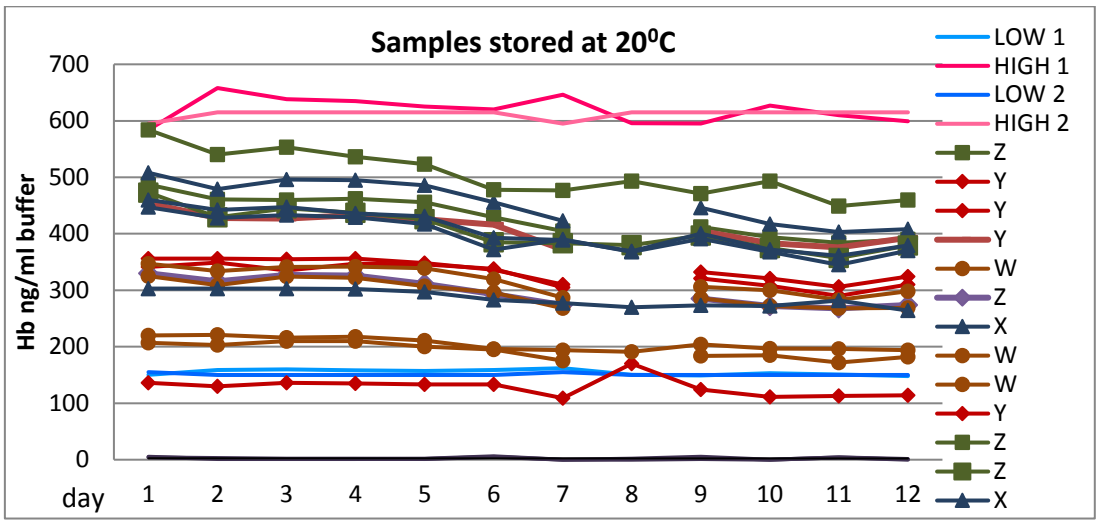
of the lysate was difficult. Reproducible sampling into the collection device also proved to be difficult since the serrated end of the sampling probe is small. Due to these factors, the initial Hb concentrations show variation within the four spiked groups. These samples were stored in the collection tubes at ~ 4^0C, ~ 20^0C, and ~ 26^0C. These were re-assayed daily, for 12 days.

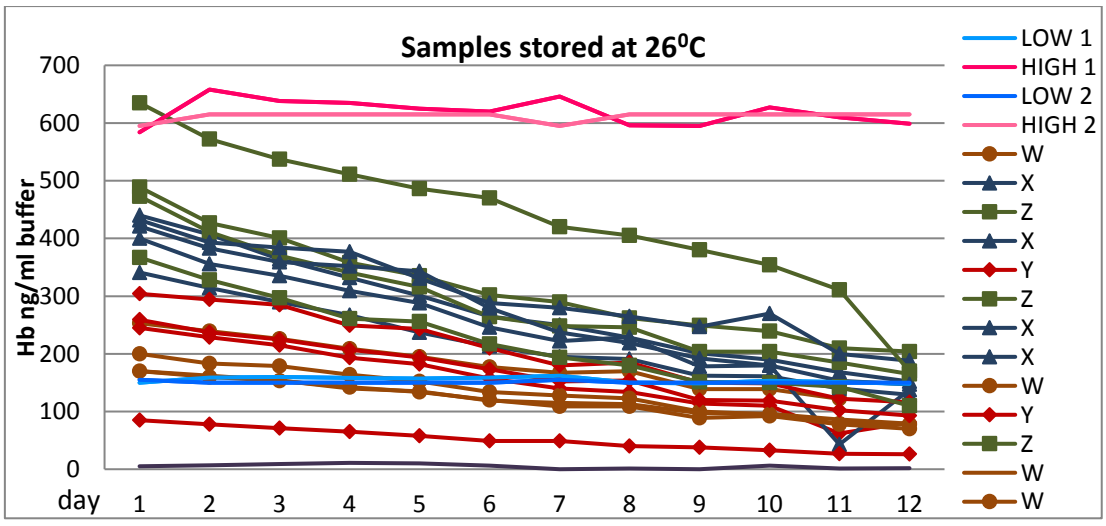
The initial concentrations of Hb in the spiked faecal samples were in the range 85–635 ng Hb/ml buffer. Five samples with undetectable f-Hb were included in the study. The results from the study showed that Hb concentration fell in all groups of samples over 12 days, as documented in Figure 14.

The rate of decrease was slightly different in each sample. This may be in part due to the difficulties in preparing the samples and/or different faecal flora present in each sample device. Overall, lower Hb was detected at each analysis as temperature increased. In samples stored at 4^0C there was a 16% decrease in f-Hb concentration, 31% at 20^0C and 61% at 26^0C. The four spiked samples were denoted as W, X, Y and Z, with W containing high Hb, X medium, Y the least and Z the highest of the group. The percentages of Hb degradation according to initial concentration are 32%, 40%, 36% and 37% respectively. The study indicated that the rate of decrease was not proportional to the original f-Hb.

Figure 14 Hb stability study data at 4, 20 and 26 ^{o}C





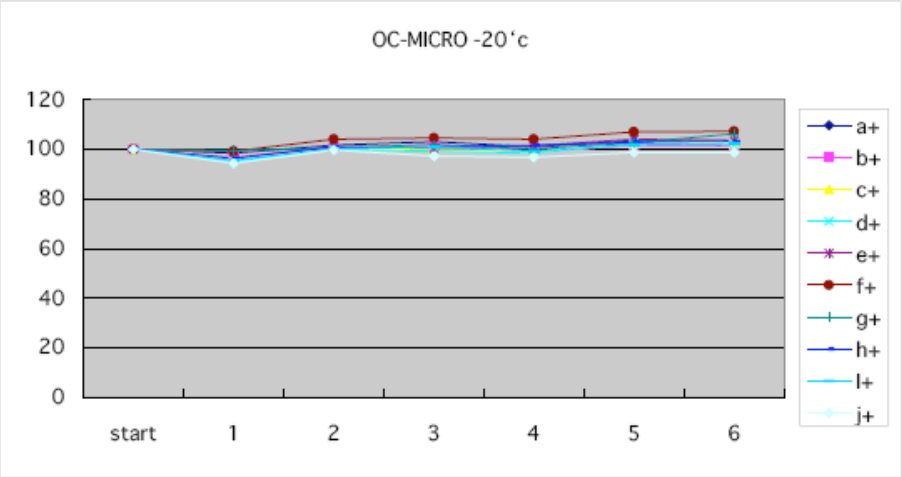


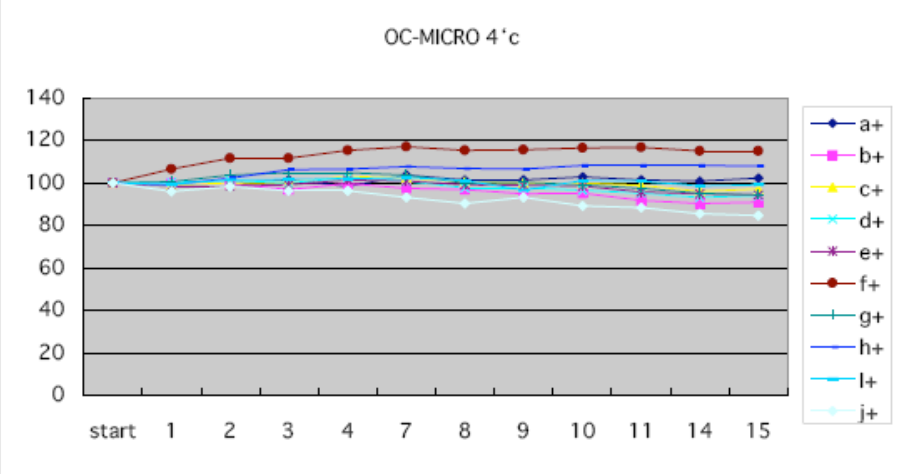
This study demonstrated that the concentration of Hb detected by automated immunochemical testing using OC-Sensor sample collection devices does decrease over time. The results also show that higher storage temperatures produce a greater reduction in the f-Hb detected. This is in contrast to gFOBT which does not suffer from this issue once the faeces has dried out (80). The rate of reduction seen was not proportional to the original f-Hb.

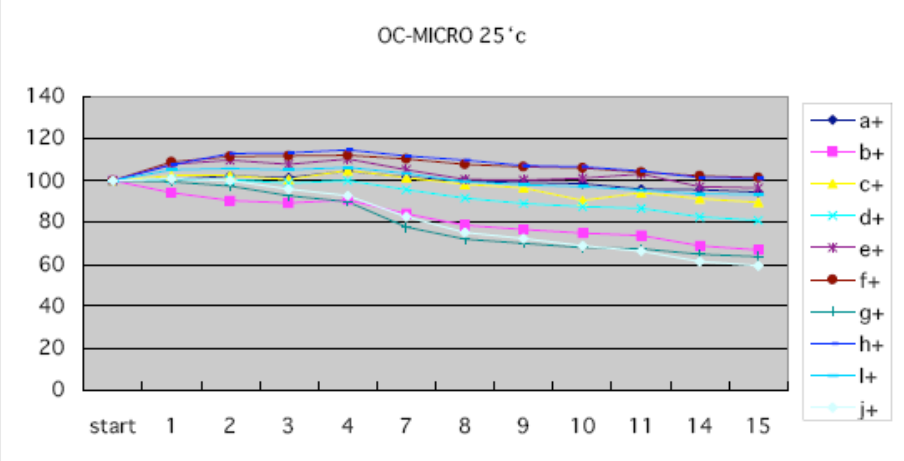
The outcomes from this work informed the decision to set the maximum time between sample collection and analysis as 10 days where temperatures tend to be under 20^0C: after this it was considered that the result would not be accurate since f-Hb may have degraded too far to give a meaningful result.

The data presented here were used to extend sample acceptance beyond the manufacturer's claims of 3 days at room temperature, 7 days at 2-10°C, 10/14 days at –20°C. These data are shown in Figure 15. Despite the desire to further extend acceptance, the decision taken must take into consideration the modality of returning samples and not attempt to include as many returned samples as possible which may not provide a meaningful result.

Figure 15 Eiken Hb stability data







3.4 Results and Reporting

The OC-Sensor analyser requires a complimentary specimen collection device. This contains 2.0 ml of buffer (± 0.13 ml) into which are suspended 10.5 mg of faeces (± 2.0 mg). The mass of faeces is standardized by use of a grooved sampling device which is passed through the neck of the buffer tube, acting as a collar to remove excess faeces, this is shown in Figure 16.

Figure 16 Diana OC-Sensor sample collection device

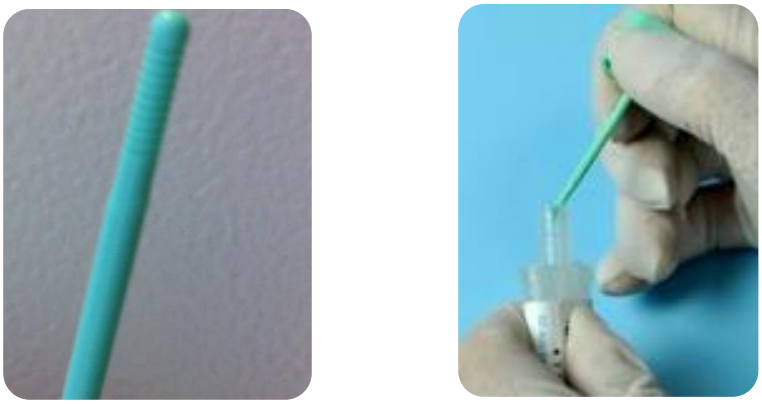


Image courtesy of the Scottish Bowel Screening Laboratory

An analysed concentration of 100 ng Hb/ml buffer corresponds to a quantity of Hb equal to 200 ng Hb in the bottle (that is in 10 mg of faeces). Thus, a conversion to µg Hb/g of faeces can be made, as shown in Table 12. These characteristics provide a reproducible way to convert ng Hb/ml buffer to µg Hb/g of faeces.

Table 12 Conversion of faecal haemoglobin concentrations (f-Hb)

ng Hb/ml buffer	µg Hb/g faeces
10	2
50	10
100	20
250	50
400	80
1000	200

The EWG on FIT for Screening of the CRCSC of the WEO recommend that units of µg Hb/g faeces be used and the conversion for this particular analyte (105). As the majority of the work undertaken in this thesis was carried out before this recommendation all results are reported as ng Hb/ml buffer.

3.5 Summary

Outcomes for the intra and inter-run imprecision in the linear range of the analyser, 40 – 1000 ng Hb/ml buffer, had a CV < 20. Where samples had f-Hb lower than the bottom of the linear range imprecision increased to CV > 20. Analysis of IQC material used on for the two analysers over the course of the study was CV <3.8% at both low (155 ng Hb/ml buffer) and high (650 ng Hb/ml buffer) control levels. The variation in results produced between the analysers was not significant. These outcomes are similar to the performance claims made by the manufacturer and were accepted as suitable for the purposes of the evaluation of FIT in the SBoSP and symptomatic study.

Measurement of the degradation of f-Hb in the sample devices at three temperatures produced results that were closely aligned to the manufacturer's data and the decision was taken to allow 10 days after the application of faeces before declaring the sample untestable.

This important, preliminary, work provided an opportunity to SBoSP team to use the OC-Sensor Diana. Staff were able to familiarize themselves with reagent usage, calibration, sample handling and minor operational issues. Despite the overwhelming evidence that automated qualitative FIT is the modality of choice for detection of FOB in faeces, it is a requirement of the ISO 15189:2012 accreditation standard to validate the manufacturer's

analytical specifications whilst an analyser is *in situ* in the laboratory and before any participant samples are tested.

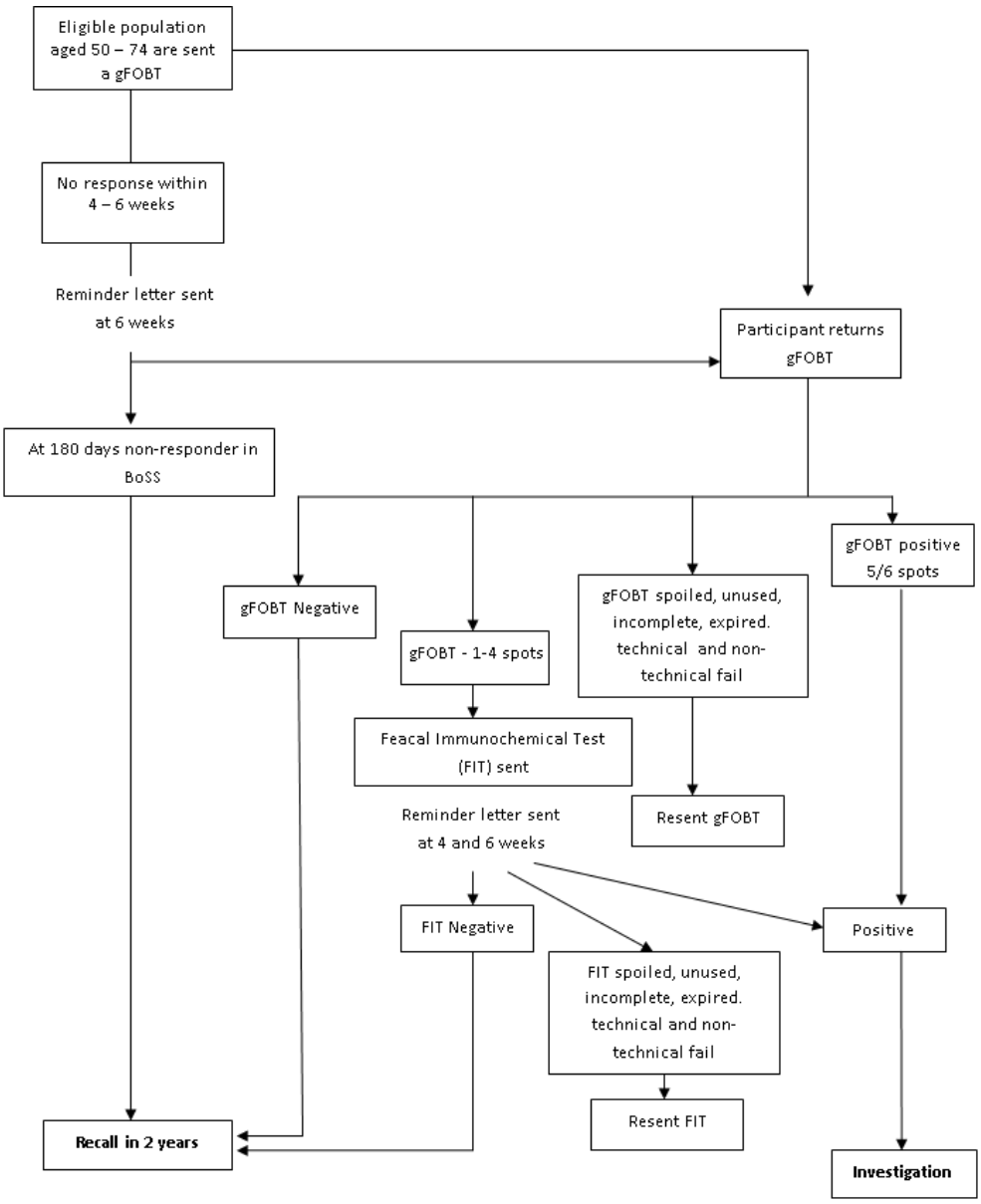
CHAPTER 4: Using FIT in Asymptomatic Population Screening

4.1 Introduction

It has been previously discussed in this thesis that, although gFOBT have advantages for use in structured screening programmes, they also have major disadvantages. The major disadvantages are poor sensitivity and specificity, and that the cut-off concentration between gFOBT negative and positive results is set by the manufacturer. Thus, the positivity and the clinical characteristics cannot be adjusted by the end-user. One way around this would be to develop a specialised algorithm locally e.g., one based on the results of initial testing with gFOBT and repeat testing with qualitative FIT as the second line test (66). This is the approach used in Scotland, resulting in a 2.1% positivity which gives the number of referrals that can be handled satisfactorily within current colonoscopy resources. However, this can tend to make programme organisation and execution complex as shown in Figure 17.

Most published studies using FIT have been performed using the cut-off as recommended by the manufacturer. This results in a higher positivity and greater sensitivity, albeit with lower specificity than a gFOBT-based programme. However, a significant consideration in Scotland, where colonoscopy capacity is limited, is that increasing the cut-off concentration reduces the positivity, in effect lowering sensitivity but with the benefit of higher specificity and fewer colonoscopies (122) (137). The use of a low referral rate for colonoscopy is necessitated by resource constraints: however, these outcomes may be relevant to other nations as healthcare resources become scarcer.

Figure 3 SBoSP algorithm, July 2010



4.2 Aims of the Evaluation

Although the suitability of FIT as a first-line test (FFLT) in Scotland could be inferred from other research studies and small scale screening programmes, there has never been a direct comparison between a fully rolled out national screening programme using gFOBT and results obtained with an automated FIT system.

The primary aim of this study was to evaluate the feasibility of using the FFLT algorithm over a six month period in two evaluation NHS Boards in Scotland with the f-Hb cut-off set to give the same positivity as seen in the SBoSP.

The secondary aims were to:

- compare the clinical outcomes with those obtained 1 year before and 1 year after in the evaluation NHS Boards,
- make a comparison with contemporaneous data in another two similar (control) NHS Boards,
- assess the effect of the single sample quantitative FIT collection device on uptake by sex, age and deprivation, as determined by Scottish Index of Multiple Deprivation (SIMD), and
- analyse clinical outcomes at 2.1% referral for colonoscopy.

4.3 Benchmarking

In Scotland, the performance of many healthcare services are reviewed through internal benchmarking. In this study the Key Performance Indicators (KPI) already being collected and published by Information Statistics Department (ISD) of NHS National Services Scotland were used as a basis for evaluation (108).

Table 13 Relevant Key Performance Indicators (KPI) for the Scottish Bowel Screening Programme (SBoSP) and expected direction of change

Benchmark	Current Outcome	Expected direction of change
1.Overall uptake of screening	53.7	↑
3.Positive screening test result rate	2.3	↔
8.Crude cancer detection rate	0.14	↑
9.Percentage of people with screen detected CRC that are Dukes' Stage A	27.7	↔
10.Percentage of people with screen detected CRC that are Dukes' Stage B	24.1	↔
11.Percentage of people with screen detected CRC that are Dukes' Stage C1	21.8	↔
12.Percentage of people with screen detected CRC that are Dukes' Stage C2	1.6	↔
13.Percentage of people with screen detected CRC that are Dukes' Stage D	2.5	↔
17.Polyp CRC detection rate	0.04	↑
18.Percentage of polyp CRC	28.0	↑
19.Overall adenoma detection rate	0.71	↑
20.High risk adenoma detection rate	0.11	↑
21.PPV to CRC	7.3	↑
22.PPV of all adenomas where adenoma is the most serious diagnosis	38.1	↑
23.PPV to high risk adenoma	5.8	↑
24.PPV to high risk adenoma or CRC	13.1	↑
25.PPV to any adenoma or CRC diagnosis	45.4	↑

Based on KPI Report Invitations between 1st Nov 2008 and 31st October 2010 (98)

Effectiveness of the screening programme is a result, *inter alia*, of participation and sensitivity of the test for neoplasia. Those KPI pertaining to measuring effectiveness were selected for review and it was hypothesised in which direction they would change, as shown in Table 13. The outcomes of the study and impact on the KPI are discussed in the Results section of this Chapter.

4.4 Methods

The study was carried out in three phases:

Phase 1 – collection of training data

Planning assumptions for the rollout of the SBoSP included positivity at 2.1%. This was based on the outcomes of the first three pilot rounds (46). The positivity in the SBoSP at the time of planning the study was 2.3% (108). Data from Dutch quantitative FIT studies (discussion with Leo Van Rossum) was used to convert this positivity back to f-Hb cut-off (118). The analyser cut-off was set at 400 ng Hb/ml buffer (equivalent to 80 µg Hb/g faeces). This would become the 'training' data and would be reviewed after one month. The resulting data would:

- reveal the actual positivity in the evaluation group compared to that in the SBoSP
- allow adjustment of the f-Hb cut-off to give a 2.1% positivity.

Phase 2 – remaining study period

At this point, if it was required, an adjustment would be made to the cut-off based on the training data. For the remainder of the study period, individuals with a test result above the revised threshold would be defined as positive.

Phase 3 - conclusion

At the end of the six month period, the NHS Boards in the evaluation reverted to using gFOBT/FIT two tier reflex algorithm. Full analysis of the evaluation data was performed at this point.

The initial planning for the SBoSP was based on Health Boards agreeing to resource colonoscopies on a 2.1% referral rate. The challenge in planning for this study was to determine a suitable f-Hb cut-off to give a positivity in the study boards that was similar to that in the programme. Dutch researchers had been very active in FIT research and van Rossum

performed analysis of FIT data from an average risk screening population tested on the OC-Sensor (137). These calculations are shown in Table 14. These data provide evidence that using one FIT sampling device with a cut-off set to ensure a similar number of referrals for colonoscopy as the current programme, should detect similar numbers of cancer and adenomas to the current gFOBT/FIT algorithm in Scotland. This data was used to inform the cut-off for the training data set in the SBoSP. The 400 ng Hb/ml buffer cut-off was retained for the entire study period.

TABLE 14 Calculation of cut-off based on data from The Netherlands (FIT is reported as ng Hb/ml buffer)

	gFBOT	FIT 50	**FIT 400**	FIT 500
Positive (N)	103	428	**122**	104
Scoped positives (%)	1.7%	7.0%	**2.0%**	1.7%
CRC	11	28	**16**	16
Advanced Adenomas	46	1.6	**53**	46
Extra CRC compared to gFOBT		2.5	**1.5**	1.5
Number needed to scope				
CRC	9.4	15.3	**7.6**	6.5
CRC + Advanced Adenomas	1.8	2.3	**1.8**	1.7

The cut-off f-Hb was designed to give a positivity that could be managed by the available colonoscopy resources. Positive screening referrals would continue at 2.4% of the population that returned the collection devices, thereby picking up those with the highest concentration of Hb in their faeces. These are the group who have the highest probability of significant neoplasia. (140).

Analyses were carried out in the SBoSL by trained staff whose major function is to perform faecal test analyses: the SBoSL has a comprehensive total quality management system and is accredited to ISO 15189 based standards by CPA (UK) Ltd. The analytical strategy and performance achieved have been detailed previously in Chapter 3. Briefly, the analysers

were calibrated once per month, or when the latex LOT changed, with the calibrators provided. Each analytical run was preceded by analysis of two quality control materials at different Hb concentrations. The target values for the lots of materials used were set a priori by 20 replicate analyses and a 1_2s rule (where the mean plus or minus two standard deviations are used as the control limits) used for acceptance or rejection of analytical runs.

Chapter 3 also reported that high temperatures and delayed testing may contribute to a reduction in the concentration of Hb detected at analysis. Because this effect is known, any samples returned more than 10 days after the date of sampling were deemed ‘expired’. Since this preliminary work was undertaken further studies have been published that examine the stability of f-Hb in sample device tubes that reinforce our findings, and endorse the use of 14 days as the point of expiry in next generation sample devices. Lower storage temperatures reduce f-Hb degradation (141) (142).

Faecal samples that have Hb concentration near the cut-off would be likely to fall below the cutoff if subjected to high temperatures and a prolonged delay in testing. From the experience gained in this study the ‘instruction for use’ and invitation letter sent to the participant extols them to return the sample device soon as possible. The protocol within the SBoSL was to test samples on the day of arrival. If this was not possible, the sample would be stored at 4^0 C for testing within 24 hours. As far a possible this reduced the potential for the concentration of f-Hb to be further reduced after arrival in the SBoSL.

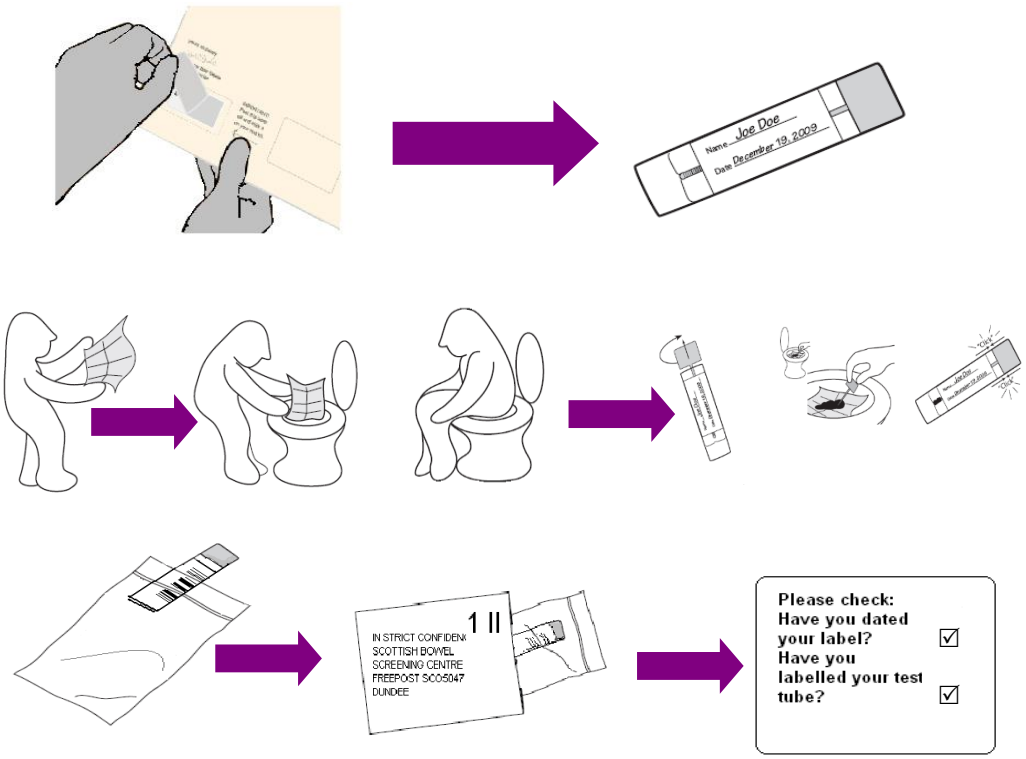
4.4.1. Population and Evaluation Period

There are 14 territorial NHS Health Boards in Scotland. For this FFLT study, all individuals in the NHS Tayside and NHS Ayrshire & Arran areas of residence from 01 July 2010 to 12 January 2011 were sent a study invitation pack. NHS Tayside was in the fifth round of screening and so prevalent CRC have been detected and mainly incident CRC emerging. NHS Ayrshire and Arran were part way through their prevalence round of screening and so it was anticipated that would be more neoplasia detected in this group in the study.

Instructions for use

All those eligible to participate in the SBoSP receive a letter of invitation, a 'know the facts' leaflet, instructions for use (IfU), a test kit, initially a gFOBT, cardboard applicators and a freepost foil return envelope.

Figure 18 Information for Users (IfU) for FIT as a First-Line Test (FFLT) Evaluation



The IfU that are sent out to eligible SBoSP participants would not be applicable to those in the study. The Senior Management Team, a multidisciplinary group including clinical and management staff, at the Scottish Bowel Screening Centre (SBoSC) developed new IfU (adapted from the manufacturers' literature) and these were sent to participants in the study (Figure 18).

Helpline activity

The SBoSC also hosts the helpline for the service. The staff are specifically trained to deal with bowel screening queries but do not offer medical advice. Training sessions were organised to pre-empt questions that might arise during the study and template answers devised to ensure equity of service. During the study calls regarding the FFLT study were monitored and recorded.

4.4.2. Invitation to the Study

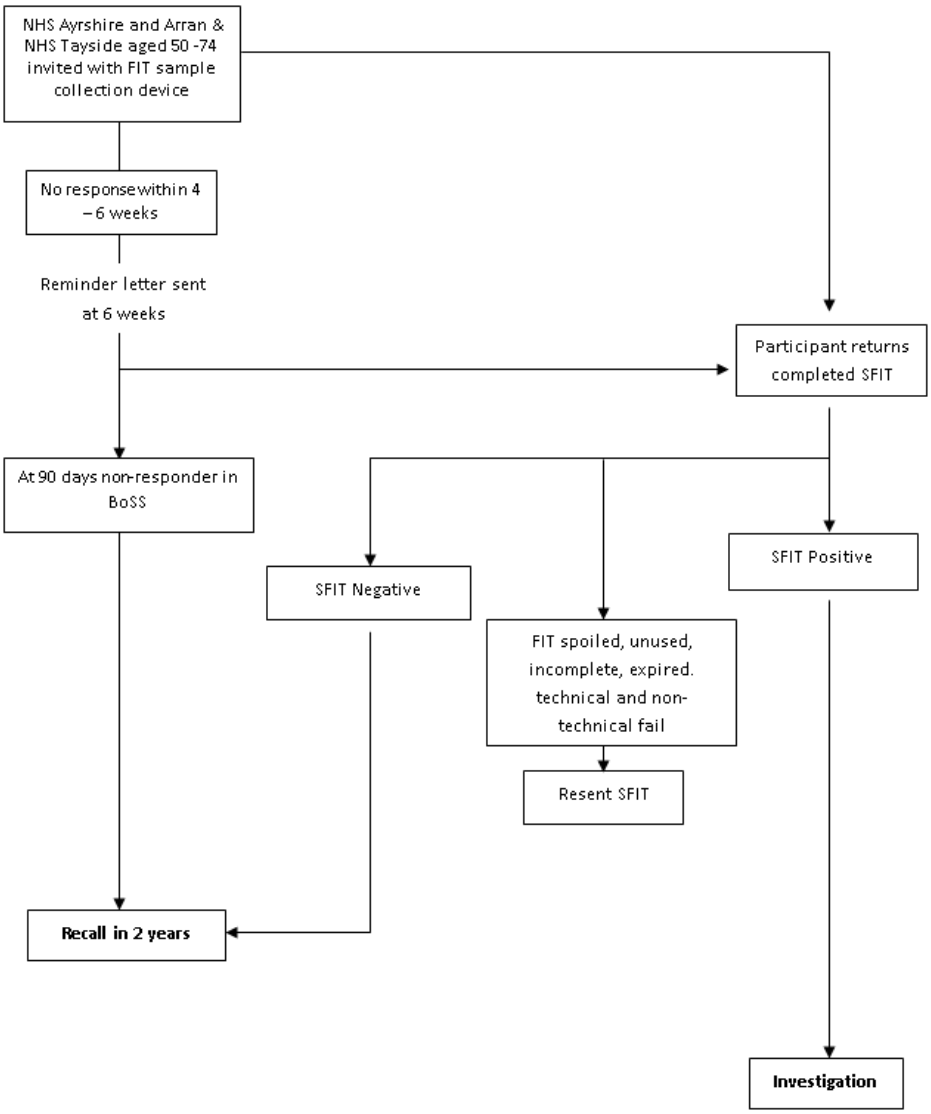
Each person in NHS Tayside and NHS Ayrshire& Arran eligible to participate in screening during the study period was sent a FFLT invitation pack. The invitation pack was based on the current SBoSP pack and contained an invitation letter, a booklet on CRC and a thin card wallet with printed written and pictorial instructions, Figure 16, for sample collection. Inside the wallet was a single faecal sample collection device, a small zip-lock plastic bag with integral absorbent material and a foil mailing pouch for device return.

The invitation letter contained an integral SBoSP identification label which was removed by the participant and attached to the outside of the zip-lock bag before return of the sample. Participants were guided through the sampling process by the pictorial instructions which had been devised specifically for the study. All study material was developed in the SBoSC

and approved by Health Scotland. Further advice about completing the sampling process could be gained by contacting the SBoSP helpline by telephone.

Invitees were allocated a compliance period of 90 days from the date of invitation to return their specimen collection device. A reminder letter was sent out at 6 weeks as per the SBoSP algorithm. Invitees who contacted the helpline to opt out or where post was undelivered were treated as per the usual practices of the SBoSP. The algorithm used for the study is set out in Figure 19.

Figure 49 SBoSP algorithm, 1 July 2011 – 12 January 2012



4.4.3. Sample Handling

The unique SBoSP label documented the name of the participant, CHI and accession number. The CHI number is a unique 10-digit identifier used ubiquitously in NHS Scotland to access healthcare; it identifies the date of birth, sex and area of residence of the individual. The label also had a 10-digit accession number generated by the BoSS IT system. The participant was required to write the date of sample collection on this label and to return the sample back to the SBoSL without delay. The foil mailing pouches, with completed sample collection devices in the zip-lock bags, were returned through the normal UK Post Office mail system by business class freepost.

On return to the SBoSL, the foil mailing pouches were opened and the label on the zip-lock bag replicated using in-house software. This secondary label was fixed to the sample collection device and the receipt of a sample captured electronically by BoSS during the 'book-in' process which gave confirmation of the name, CHI number and kit number via scanning of the barcode. Samples that were received >10 days from date of sample collection were termed "expired" and not tested further. This decision was based on in-house validation studies of Hb stability performed during the commissioning of the analysers as described in Chapter 3.

Analyses were carried out in the SBoSL by trained staff whose major function is to perform faecal test analyses. The SBoSL has a comprehensive total quality management system and is accredited to ISO 15189 based standards by CPA (UK) Ltd. The analytical strategy and performance achieved have been detailed in Chapter 3.

Once in the SBoSL 93.3% of returned samples were tested on the day of receipt; where this did not happen samples were stored at 4^{o}C. During the course of the evaluation, the maximum storage at 4^{o}C was 8 days (owing to delays shipping reagents during adverse weather). These samples were allowed to return to room temperature before analysis.

4.4.4. Reporting of Participant Results

The OC-Sensor Diana uses a Windows® based operating system controlled by a touch screen interface. It has sufficient software to construct and display calibration curves and collect control data, allowing the user to review retrospective control data and examine these for imprecision and bias. Unfortunately, the test result data can only be printed via a ticker tape set up or be imported into a comma-separated values (.csv) file. As part of the work up for using this analyser, representatives from MAST Group undertook construction of an interface programme that allowed the conversion of quantitative results into qualitative data which was produced in a format allowing results to be printed out linked to the participant's unique test request number. This gave positive/negative test results that could be printed onto a worksheet and entered manually by two members of staff into the bespoke Bowel Screening System (BoSS) IT interface each day (Appendix 2). These printed records where retained as per all patient records (136).

All participants with f-Hb <400 ng Hb/ml buffer were considered negative and sent an explanatory letter. Those who sent an untestable device were sent another FIT kit pack. All participants with a f-Hb ≥ 400 ng Hb/ml buffer were considered positive and also letter informing them of this. The relevant GP was notified and the individual referred to the NHS Board of residence for colonoscopy.

4.4.5. Study Groups

Individuals with f-Hb ≥ 400 ng Hb/ml buffer were referred for colonoscopy. Data on colonoscopy outcomes and any subsequent pathology in the participants with positive results were downloaded from the appropriate

NHS Tayside and NHS Ayrshire & Arran clinical IT systems for colonoscopy and pathology (Group 1).

Data on clinical outcomes for NHS Fife and NHS Forth Valley were provided by Information Services, NHS National Services Scotland (Group 2). Data on clinical outcomes for an identical period of time, at the same time of year, *before* the study (01 July 2009 to 12 January 2010) were collected as described above for NHS Tayside and NHS Ayrshire & Arran (Group 3). Similarly, data on clinical outcomes for an identical period, immediately *after* the study, (13 January 2011 to 27 July 2011) were collected for NHS Tayside and NHS Ayrshire & Arran (Group 4). The makeup of the groups is shown in Table 15.

Table 15 Identification of Groups used for analysis of clinical outcomes (in chronological order) showing time periods, NHS Boards and screening algorithms in use

01 July 2009 - 12 January 2010
Group 3 • NHS Tayside – three pilot rounds and into 2nd round of SBoSP • NHS Ayrshire & Arran - 2 months into prevalence round of screening, 4½ months 1st incidence round of SBoSP • gFOBT/FIT two-tier reflex algorithm
01 July 2010 – 12 January 2011
Group 1* • NHS Tayside – three pilot rounds and into 2nd round of SBoSP • NHS Ayrshire & Arran – 1st incidence round of SBoSP • **FFLT algorithm**
Group 2 • NHS Fife – three pilot rounds and into 2nd round of SBoSP • NHS Forth Valley – 1st incidence round of SBoSP • gFOBT/FIT two-tier reflex algorithm
13 January 2011 - 27 July 2011
Group 4 • NHS Tayside - three pilot rounds and into 2nd/3rd round of SBoSP • NHS Ayrshire & Arran - 1st incidence round of SBoSP • gFOBT/FIT two-tier reflex algorithm

* Evaluation Group

4.4.6. Statistical Analysis

Data on sex and age were determined from the CHI number. Number, size and location of CRC and adenomas were recorded. Assignment to outcome groups was as recommended by the British Society of Gastroenterology (BSG) (48), but also in accordance with the protocol for patient follow-up used in Scotland in that the higher (high and Intermediate) risk adenoma (HRA) group was based on combining the intermediate and high-risk groups identified by the BSG.

Assuming a 60% uptake, the participants in this study were expected to generate 840 colonoscopies at 2.1% positivity. With the current algorithm, this would be expected to detect about 80 CRC and 300 adenoma cases. Overall, these numbers were sufficient to detect, with 85% power, an increase in uptake from 55% to 60%, an increase in CRC detection rate from 0.07% to 0.17% and an increase in PPV from 7.5% to 10.0%.

Statistical analysis of uptake was undertaken by comparing those invited during the FIT study period and the corresponding six month time periods from previous and subsequent years in the comparator NHS Boards. Chi-squared tests were used to assess the change in uptake during the FIT study period compared with uptake in the other time periods combined and weighted. MedCalc (MedCalc Software, Mariakerke, Belgium) statistical software was used for all calculations. Probability of $p<0.05$ was considered significant.

4.5 Results

The results recorded reflected the programme KPI's discussed earlier in the Chapter.

4.5.1 Number of Invitations Sent Out and Sample Devices Returned

Study kits were mailed to participants from 1 July 2010 and ceased to be sent to participants on 12 January 2011. The number of packs [flow-pack with sample collection device, return envelope and zip-lock bag, IfU and the Know the Facts leaflet] sent out from the SBoSC as initial invitations was 66,225 of those 40,125 were returned.

Uptake

Uptake, defined as the percentage of invitees (66,225) who completed their cycle with a positive or negative test result (38,720), was 58.5%. In 99.2% of cases the sample was analysed three or fewer days after receipt of sample: 93.3% were analysed on the day of receipt.

Before and after the study period uptake in the four NHS Boards was recorded for periods of six calendar months. Data for the Study Boards has been truncated to 31 December 2010 to match the rest of the data presented. This gave uptake for three different periods as shown in Table 16. The p-values demonstrate that the uptake was significantly higher during the FFLT evaluation period in NHS Tayside and NHS Ayrshire & Arran than during the time periods before and after.

Table 16 Uptake (%) in four NHS Boards for three six-month periods. (FIT Study group in bold.)

	Tayside	Ayrshire & Arran	Fife	Forth Valley
01 Jul 2009 – 31 Dec 2009				
Invited	37275	31713	30685	21972
Accepted (%)	20764 (55.7)	16491 (52.0)	16311 (53.2)	11640 (53.0)
01 Jul 2010 – 31 Dec 2010				
Invited	**32195**	**30570**	29397	22881
Accepted (%)	**19600 (60.9)**	**17742 (58.0)**	15425 (52.5)	11626 (50.8)
01 Jul 2011 - 31 Dec 2011				
Invited	37153	32450	30938	22898
Accepted (%)	20274 (54.6)	16790 (51.7)	16044 (51.9)	11848 (51.7)
***p*-value**	<0.0001	<0.0001	0.102	0.036

NHS Tayside and NHS Fife were compared since both NHS Boards participated in the screening pilots, they are of a similar size and have similar population demographics. NHS Ayrshire & Arran and NHS Forth Valley did not participate in the early pilot rounds of screening and have similar populations (144). Figure 20 and 21 show comparisons of uptake in each of the two NHS Board groups.

Figure 20 Comparison of uptake (%) in NHS Tayside and NHS Fife before, during and after the study period

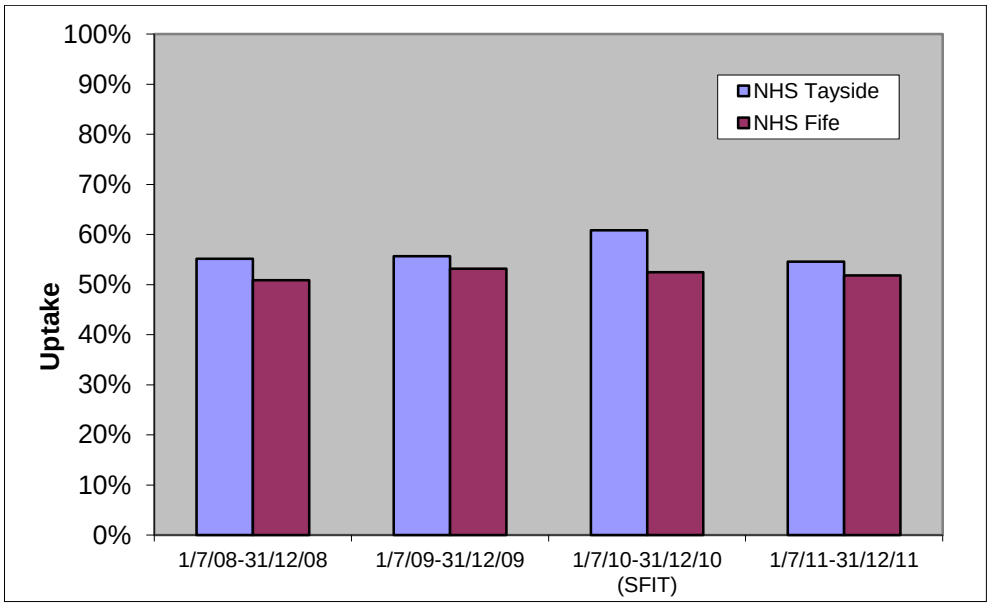
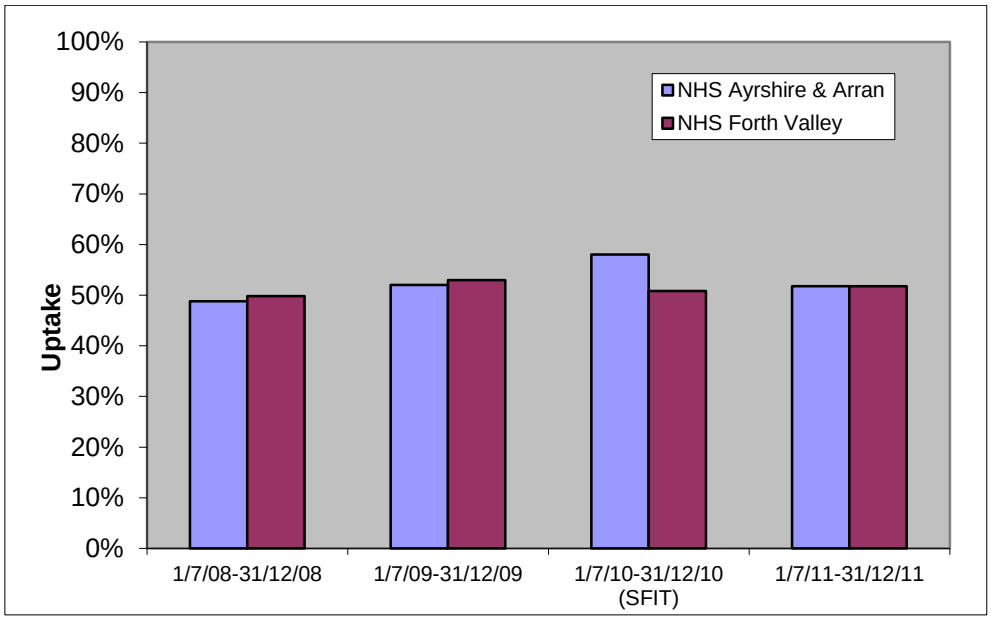


Figure 21 Comparison of uptake (%) in NHS Ayrshire & Arran and NHS Forth Valley before, during and after the study period



This is the first time that it has been shown that uptake rates fall back to previous levels with the reintroduction of the gFOBT/FIT algorithm.

Untestable Samples

Participants who return samples to the SBoSL may return samples that are untestable. These fall into the following categories: spoiled, expired, unused, incomplete and technical fail. The return rate for repeat samples was more than 80% in each category, as shown in Table 17. This is higher than the uptake of the initial screening tests, irrespective of whether gFOBT or quantitative FIT.

Experience in the SBoSP indicates that the return rate of the reflex qualitative FIT is approximately 95%, which again is higher than uptake for the initial screening test.

Table 17 Return of repeat ‘untestable’ sample devices

Expired			**Non-Technical Fail**			**Spoiled**		
Condition	**Number**	**%**	**Condition**	**Number**	**%**	**Condition**	**Number**	**%**
Returned	1134	86.1	Returned	34	94.0	Returned	41	83.0
Not returned	171	13.0	Not returned	2	5.6	Not returned	8	16.3
Total	1305		Total	36		Total	49	
Negative	1051	80.2	Negative	34	94.4	Negative	31	63.3
Positive	32	2.4	Positive	0	0.0	Positive	3	6.1
Expired	40	3.1	Expired	0	0.0	Expired	1	2.0
Spoiled	4	0.3	Spoiled	0	0.0	Spoiled	6	12.2
Closed	167	12.7	Closed	2	5.6	Closed	8	16.3
Returned	0	0.0	Returned	0	0.0	Returned	0	0.0
Not Returned	3	0.2	Not Returned	0	0.0	Not Returned	0	0.0
Others	8	0.6	Others	0	0.0	Others	0	0.0
Total	1305		Total	36		Total	49	

Further Outcomes

A major rationale for the study was to evaluate whether FFLT could be delivered within the context of the SBoSL. Aspects of service delivery that were tested during the study were staff training, set-up and maintenance of analysers and turnaround times for results. Overall, the planning and delivery of the evaluation went smoothly with no major problems. Over 99% of results were reported within three days of receipt in the Laboratory.

The vast majority of participants received an unequivocal result in less than two weeks from sample collection. 98% of SFIT were returned within 10 days of the sampling date. This would have been higher if not for the extreme adverse weather in November 2010. A detailed report regarding all laboratory outcomes is included in Appendix 3.

There were very few calls to the Helpline regarding SFIT and its use: 205 calls from the 66 225 invitations issued (0.3% compared to 6% with the current gFOBT/FIT algorithm). Focus group work found that the pictorial and written instructions for use were very clear. Participants seemed to find the

single stool sampling method and sampling process straightforward. The analysis of helpline activity is shown in Appendix 4.

4.5.2 Characteristics of Those Who Returned a Sample Device

The number and percentages of participants with a positive test result stratified into 5-year age groups are shown in Table 18 for Groups 1, 3 and 4: similar data for Group 2 were unavailable. As expected, the proportion of participants receiving a positive test result increased with increasing age and was higher in males than in females when the non-stratified totals are compared.

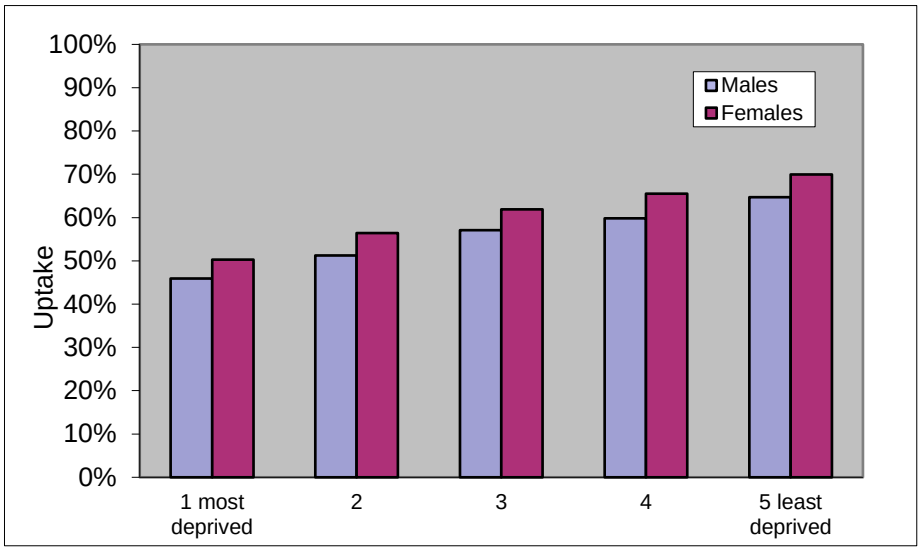
Table 18 Number (%) of participants with a positive test result by sex and age. * indicates the evaluation Group

	Group 1*	Group 2	Group 3	Group 4
Total	943	736	732	626
Males	532 (56.4)	401 (54.5)	443 (60.5)	376 (60.1)
Females	411 (45.6)	335 (45.5)	289 (39.5)	250 (39.9)
50-54 years	151 (16.0)	n/a	99 (13.5)	88 (14.1)
55-59 years	189 (20.0)	n/a	139 (19.0)	129 (20.6)
60-64 years	173 (18.3)	n/a	131 (17.9)	129 (20.6)
65-69 years	206 (21.8)	n/a	168 (23.0)	136 (21.7)
70-74 years	224 (23.6)	n/a	195 (26.6)	144 (23.0)

n/a data were unavailable from Information Services Division, NHS National Services Scotland

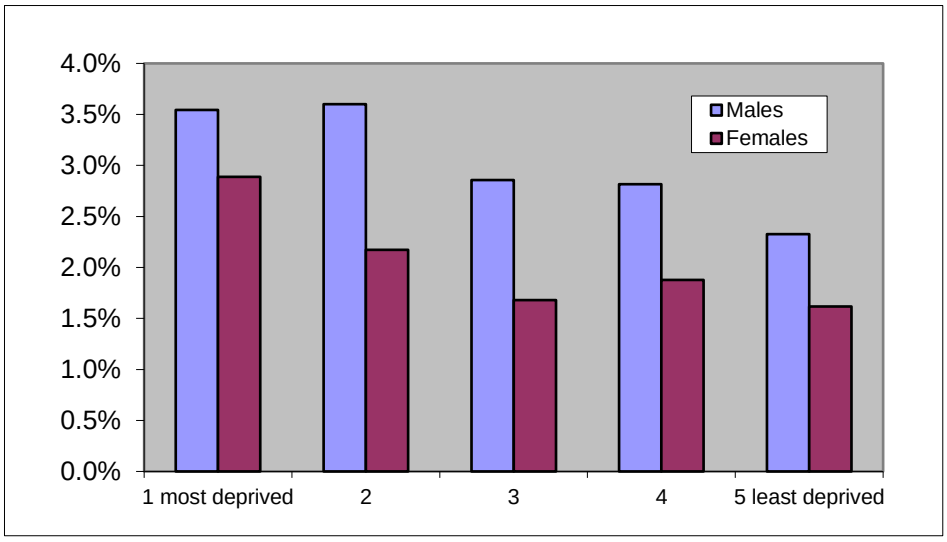
The effects of deprivation and sex on Group 1 are illustrated below, Figure 22. As expected from SBoSP data, uptake fell with increasing deprivation and was greater in women than in men. Uptake ranged from 45.8% in the most deprived males up to 69.9% in the least deprived females.

Figure 5 Uptake and Scottish Index of Multiple Deprivation (SIMD) quintile of Group 1



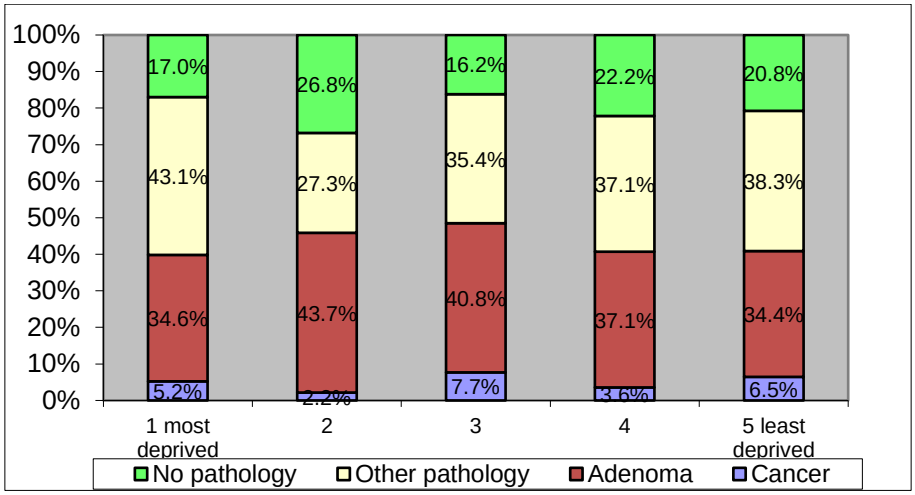
A higher percentage of males had a positive result than females in every deprivation quintile and positivity was at its lowest in the least deprived and highest in the most deprived as shown in Figure 23. The least deprived females, therefore, had the lowest positivity, at 1.6%, which was less than half of that of the highest positivity which was found amongst the most deprived males, at 3.5%.

Figure 23 SIMD Quintile of participants who received a positive test result



Interestingly, however, there was no effect of deprivation on pathology, Figure 24, whereas a negative association with increasing deprivation has been demonstrated with gFOBT (145). It can be seen that uptake was higher in both sexes and all deprivation categories than has been observed with gFOBT in either of the evaluation NHS Boards.

Figure 24 SIMD quintile showing colonoscopy outcome in evaluation group



4.5.3 Clinical Outcomes of Those with a Positive Test Result

A complete data set was collected for those patients with a positive test result and this is shown in Table 19. This demonstrates that the clinical outcomes of those who participated in the evaluation (Group 1) was essentially the same as in the three other Groups. It is worth noting that the PPV for CRC was lower for Group 1 participants (4.8%) than Group 3 participants (7.7%) – p = 0.0291 – this is because in newly screened populations areas there is a higher prevalence of disease than in a screened and treated population. Screening followed by action reduces the prevalence of the disease in the target population.

There were 943 participants with a positive test result in Group 1**,** 453 (48.0%) from participants in NHS Tayside and 490 (52.0%) from participants in NHS Ayrshire & Arran, For Group 2, there were 736 positive results, 383 (52.0%) from NHS Fife and 353 (48.0%) from participants in NHS Forth Valley. For Group 3, there were 732 positive results, 374 (51.1%) from NHS Tayside and 358 (48.9%) participants from NHS Ayrshire & Arran. For Group 4, there were 626 positive results, 280 (44.7%) from participants in NHS Tayside and 346 (55.3%) from participants in NHS Ayrshire & Arran in the prevalent screening round.

Essentially, the quantitative FIT performed as expected when taking into account the effect of participants completing various rounds screening in the Groups. The prevalent screening round encompasses 100% of individuals who have never undertaken screening. Incident rounds are the subsequent screening rounds: in such rounds a small number, 50 year olds and others new to screening will be in a prevalent round for them.

Table 19 Clinical outcomes in participants with a positive test result

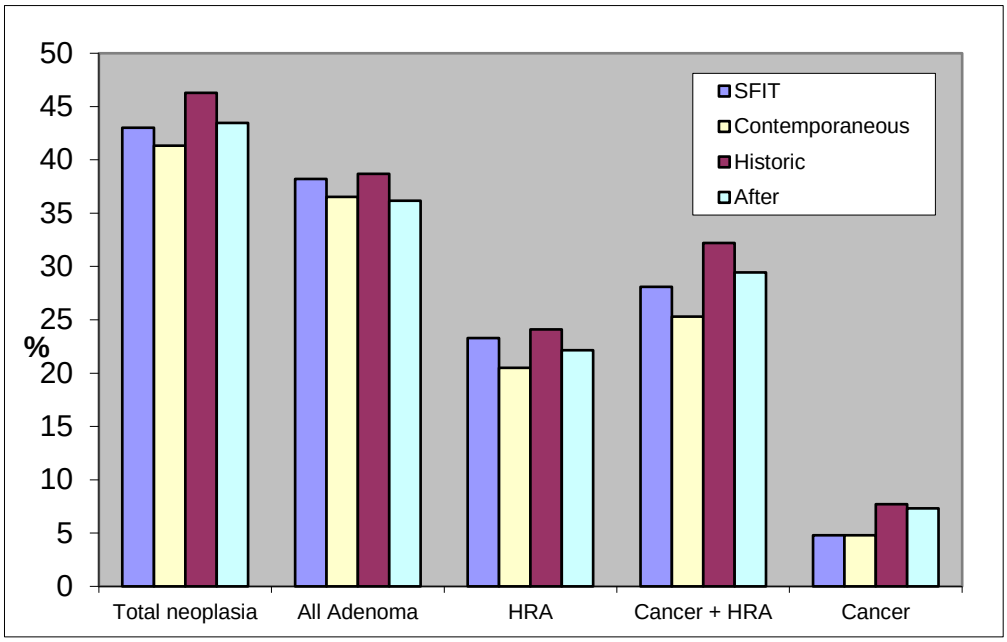
	Group 1						**Group 2**					
	Total		**Males**		**Females**		**Total**		**Males**		**Females**	
	n	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
Participants with positive result	943		532		411		736		401		335	
No investigations /incomplete investigations/outcome unknown/excluded	129	(13.7)	73	(13.7)	56	(13.6)	130	(17.7)	65	(16.2)	65	(19.4)
Investigations completed	814	(86.7)	459	(86.6)	355	(86.8)	606	(82.3)	336	(83.8)	270	(80.6)
Clinical Outcomes	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**
CRC	39	4.8	23	5.0	16	4.5	33	5.4	19	5.7	14	5.2
High-risk adenoma (HRA)	190	23.3	127	27.7	63	17.7	115	19.0	80	23.8	35	13.0
CRC + HRA	229	28.1	150	32.7	79	22.3	148	24.4	99	29.5	49	18.1
All adenoma	311	38.2	205	44.7	106	29.9	217	35.8	139	41.4	78	28.9
Total neoplasia (CRC + all adenoma)	350	43.0	228	49.7	122	34.4	250	41.3	158	47.0	92	34.1
Hyperplastic polyps	64	7.9	40	8.7	24	6.8	-	-	-	-	-	-
Normal/Other pathology (IBD, DD, angiodysplasia, haemorrhoids etc,)	400	49.1	191	41.6	209	58.9	-	-	-	-	-	-

	Group 3						Group 4					
	Total		Males		Females		Total		Males		Females	
	n	%	n	%	n	%	n	%	n	%	n	%
Participants with positive result	732		443		289		626		376		250	
No investigations /incomplete investigations/outcome unknown/excluded	80	(10.9)	45	(10.2)	35	(12.1)	84	(13.4)	58	(15.4)	26	(10.4)
Investigations completed	652	(89.8)	398	(90.5)	254	(88.8)	542	(86.6)	318	(84.6)	224	(89.6)
Clinical Outcomes	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**	**n**	**PPV**
CRC	50	7.7	35	8.8	15	5.9	38	7.0	26	8.2	12	5.4
High-risk adenoma (HRA)	157	24.1	114	28.6	43	16.9	120	22.1	88	27.7	32	14.3
CRC + HRA	207	31.7	149	37.4	58	22.8	158	29.2	114	35.8	44	19.6
All adenoma	252	38.7	181	45.5	71	28.0	190	35.1	130	40.9	60	26.8
Total neoplasia (CRC + all adenoma)	302	46.3	216	54.3	86	33.9	228	42.1	156	49.1	72	32.1
Hyperplastic polyps	58	8.9	40	10.1	18	7.1	32	5.9	19	6.0	13	5.8
Normal/other pathology (IBD, DD, angiodysplasia, haemorrhoids etc,)	292	44.8	142	35.7	160	63.0	284	52.4	143	45.0	141	62.9

- these data were unavailable from Information Services Division, NHS National Services Scotland

This affects the amount of disease detected, Figure 25 shows disease detection was greatest in the historic NHS Tayside and NHS Ayrshire & Arran gFOBT/FIT group and least in the same NHS Boards after the study period. The Group 1 lie between these two groups in chronology and magnitude.

Figure 25 Percentage of participants in each clinical outcome group



4.5.4 Clinical Outcomes at 2.1% Positivity

Having complete data for the 2.4% of Group 1 referred for colonoscopy means these are available to analyse the outcomes if the positivity was set at 2.1% as per the SBoSP planning assumptions. In this scenario, 775 of the 943 actually referred would have positive results. Taking the 775 participants with the highest f-Hb in the FFLT evaluation and looking at F-Hb would mean changing the f-Hb cut-off to 526 ng Hb/ml buffer. Clinical outcomes of the 168 participants who were positive in the FFLT evaluation with f-Hb above 400 ng Hb/ml buffer but below 526 ng Hb/ml buffer, are shown in Table 20. Table 21 shows the clinical outcomes of the 775 participants who would be deemed positive using the 526 ng Hb/ml buffer cut-off and Table 22 shows, for comparison, the overall outcomes seen with the actual Group 1 participants with the cut-off at 400 ng Hb/ml buffer (positivity: 2.4%).

Table 20 Outcomes where f-Hb is 400 – 526 ng Hb/ml buffer

	Males		Females		**Both sexes**	
	n	%	n	%	**n**	**%**
Positive test result	90		78		**168**	
No investigations/ Outcome unknown /Excluded	5	5.6	12	15.4	**17**	**13.7**
Investigations completed	85	94.4	66	84.6	**151**	**89.9**
Clinical Outcomes						
Total neoplasia	41	48.2	21	31.8	**62**	**41.1**
All adenoma	38	44.7	21	31.8	**59**	**39.1**
Higher risk adenoma (HRA)	17	20.0	6	7.7	**23**	**15.2**
CRC + HRA	20	23.5	6	7.7	**26**	**17.2**
CRC	3	3.5	0	0	**3**	**2.0**

Table 21 Outcomes where f-Hb is ≥ 526 ng Hb/ml buffer

	Males		Females		**Both sexes**	
	n	%	n	%	**n**	**%**
Positive test result	442		333		**775**	
No investigations/ Outcome unknown/Excluded	68	15.4	44	13.2	**112**	**14.5**
Investigations completed	374	84.6	289	86.8	**663**	**85.5**
Clinical Outcomes						
Total neoplasia	187	50.0	101	34.9	**288**	**43.4**
All adenoma	167	44.7	85	29.4	**252**	**38.0**
Higher risk adenoma (HRA)	110	29.4	57	19.7	**167**	**25.2**
CRC + HRA	130	34.8	73	25.3	**203**	**30.6**
CRC	20	5.3	16	5.5	**36**	**5.4**

Table 22 Outcomes where f-Hb is ≥ 400 ng Hb/ml buffer

	Males		Females		**Both sexes**	
	n	%	n	%	**n**	**%**
Positive test result	532		411		**943**	
No investigations/ Outcome unknown/Excluded	73	13.7	56	13.6	**129**	**13.7**
Investigations completed	459	86.3	355	86.4	**814**	**86.3**
Clinical Outcomes						
Total neoplasia	228	49.7	122	34.4	**350**	**43.0**
All adenoma	205	44.7	106	29.9	**311**	**38.2**
Higher risk adenoma (HRA)	127	27.7	63	17.7	**190**	**23.3**
CRC + HRA	150	32.7	79	22.3	**229**	**28.7**
CRC	23	5.0	16	4.5	**39**	**4.8**

PPV for CRC and HRA are lower in the group with f-Hb 400 – 526 ng Hb/ml buffer than in the group with f-Hb above 526 ng Hb/ml buffer that would give the 2.1% positivity (2.0% and 15.2% compared with 4.8% and 23.3%, respectively), although raising the f-Hb cut-off would mean three CRC and

23 HRA would have been missed in this group. The three CRC that would not have been detected were one rectal cancer with Dukes' stage C1, one Dukes' B rectal cancer and one 7 mm diameter polyp CRC located in the sigmoid colon.

4.6 Discussion and Conclusions

The overall uptake of the FIT (58.5%) was greater than that achieved throughout the three pilot screening rounds undertaken in Scotland, which used gFOBT only (46). The KPIs for the SBoSP, which uses the two-tier gFOBT/FIT algorithm, provide benchmarking for uptake and clinical outcomes of screened individuals. The data available for the period 01 November 2008 to 31 October 2010 show uptake was 53.7% overall in Scotland and 53.1% in NHS Ayrshire & Arran and 58% in NHS Tayside (108). Uptake during the FFLT study rose by 5.2% and 6.0% in the NHS Boards, respectively. Attainment of this increase meets one of the goals of the evaluation in assessing introduction of a new test modality on the effectiveness of a fully rolled out SBoSP. In addition, the data gathered showed that, when the study ended, uptake fell to similar values to those previously seen with the current algorithm. The two NHS Boards in which FIT was not used had small changes in uptake over time but did not have the important rise in participation seen with use of the quantitative FIT.

It is not surprising that uptake with a single sample collection device is higher than that in the SBoSP which uses gFOBT as an initial test since gFOBT require two samples from each of three faeces and, although, those with five or six windows positive are referred directly for colonoscopy, whereas those with one to four windows positive are required to undertake another test. This finding is similar to the results found in other studies comparing uptake with gFOBT and FIT (146).

Development of the study materials necessitated making changes to the information sent to participants. This presented an opportunity to develop a different style of communication and a pictogram was devised to be displayed on the A5 sized wallet containing the sample collection device. This new packaging style may have seemed modern and appealing and may well have contributed to the increase in uptake. The number of participants who were unable to complete their test kit was lower in the study group than in the SBoSP as were the number of calls to the helpline. Unfortunately, it was not possible to evaluate the psycho-social aspects of the information for users during the study: however, these very positive findings provide support for the widely-held view that more user-friendly faecal collection devices and sampling regimes encourage participation in screening programmes. Further analysis of sample return rates showed a significant increase in uptake in the most deprived groups and particularly in deprived males. Although, Symonds et al state that 'It is concerning that the lowest participation rates for FIT screening occur in males and those who are more socio-economically deprived', increased uptake in these groups may be considered a benefit from moving to FIT (141).

In line with previous findings in bowel screening pilots and programmes using faecal tests, males accounted for a higher percentage of positive test results than females in all four Groups. The percentages of positive test results in each five year age group for Groups 1, 3 and 4 are shown in 17. In the three study groups, the highest proportion of positive results was in the 70-74 years of age quintile and, in general, positivity increased with age in both sexes. It is also known that incidence of CRC is affected by sex and age and the potential consequences of these relationships have been published in detail (1) (2). It may be that different f-Hb cut-off or different screening intervals should be applied for the different sexes and ages: this is discussed in more detail in Chapter 5.

Since this evaluation was performed in the context of a fully rolled-out operational screening programme, those participants who had results below the chosen f-Hb cut-off (400 ng Hb/ml buffer) were not investigated further. This is in contrast to the two randomised controlled trials performed in The Netherlands (118) (147) (148) and in some other studies (149) (150). Such studies have shown that, at the lower f-Hb cut-off used, positivity for FIT were higher than for gFOBT and, when reported, sensitivity was higher for FIT than gFOBT, although the specificity was lower. Thus, the gain in disease detected was offset by the number of false positive test results. It has been well documented that the sensitivity increases and the specificity decreases as the f-Hb cut-off concentration is lowered: the gain is mainly in detection of advanced neoplasia (122) (137) (147).

The PPV of the screening strategy is a vital characteristic of any programme. PPV has been shown to increase and positivity decreases as the cut-off f-Hb is increased (148) (149). The results in Table 18 demonstrate that there is very little difference between the four Groups and that the pathology found with the evaluation Group was similar to that achieved with the gFOBT/FIT two-tier reflex screening algorithm.

This finding was not unexpected since the f-Hb cut-off selected was designed to give the same positivity for both approaches. If the benefits of FIT for detection of adenomas in particular (120) are to be achieved, then a lower cut-off would be required, a greater colonoscopy resource would have to be available and the programme would have to be prepared to deal with lower specificity with a greater number of false positive test results, as shown by the decrease in PPV that occurs as f-Hb cut-off is lowered (151).

Progression through rounds of screening should reduce the likelihood of finding disease as it is detected and treated, round by round. Therefore, the differences in stage of screening between the Groups means that comparison of PPV between the Groups comes with some caveats. Group

3 included some NHS Ayrshire & Arran participants in prevalence screening and, upon further examination of this group, the PPV was found to be statistically significantly higher ($p < 0.05$) for those in the two months in prevalence screening than those in the 4½ months of incidence screening for CRC + HRA (50.0% and 31.4%), HRA (41.7% and 23.5%), and total neoplasia (60.4% and 53.6%).

Those invited within the prevalence round accounted for 20.6% of the NHS Ayrshire & Arran participants in Group 3 but had 35.0% of the CRC and 34.2% of the HRA detected. This may explain the higher PPV for all clinical outcomes in Group 3 compared to Group 1 participants: however, this does not explain the high PPV for CRC in Group 4. It is of note that Group 1 compares favourably with regard to adenoma detection and Groups 3 and 4 with regard to normal/other pathology outcomes. These data show that the clinical outcomes using quantitative FIT with a high f-Hb cut-off are similar to those gained with gFOBT and gFOBT/FIT approaches.

One of the constraints of the study is that it was performed in 2 of the 14 NHS Boards in Scotland. Data presented in this Chapter suggest that screening round affects clinical outcomes, but there is the further consideration of social and geographical factors, particularly deprivation as determined by postcode and Scottish Index of Multiple Deprivation (SIMD). Across Scotland there is variation in SIMD, and the NHS Boards participating in this study may be considered to represent a general picture (144). However there is a variation in the SIMD, diet and health choices in each board area. Geographical variation has been reported in a comparison of f-Hb in Scotland, Taiwan and Italy (151).

Chaing et al. (152) investigated whether there are differences in the performance characteristics of different analytical platforms. Data from the

Taiwanese Nationwide CRC Screening and Taiwan Cancer Registry was used to evaluate the short-term and long-term performance characteristics of two FIT analysers used in a screening setting, OC-Sensor (Eiken Chemical Co, Tokyo, Japan) and HM-Jack (Kyowa Medex Co Ltd, Tokyo, Japan).The f-Hb cut-off were set to give a positive result at 20 µg Hb/g faeces. When measuring the short term indicators, including, PPV for CRC, cancer detection rate and the interval cancer rate the OC-Sensor performed better than the HM-Jack. PPV was 6.8% and 5.2% in each analytical system and cancer detection rate was 0.21% and 0.17% respectively. The interval cancer rate, was 30.7 v 40.6 per 100,000 person years.

The study used data from 2004 to 2009, this provided an opportunity to calculate mortality, a long term indicator of the effectiveness of a screening programme. The hazard ratio for crude rates was 1.21 (95% CI, 0.91 – 1.61) for the whole cohort and reduction in mortality of 11% (95% CI, 6 – 16%) when comparing participants and nonparticipants. These outcomes were not significantly different between the two analysers. The difference in outcomes in the short-term and lack of difference in long-term indicators is not what would be expected and it may be that given more time any difference in long term outcomes will increase in magnitude.

Another of the limits of the evaluation of FFLT in the context of the SBoSP was that data were not available to determine the effects of variables other than age, gender and deprivation on screening outcomes. A recent retrospective analysis of screening and surveillance FIT results by Symonds et al. (141) found that f-Hb was higher in men and increased with age and deprivation. In addition, this group also investigated the effect of previous screening participation, seasonal variation in ambient temperature and time from sample to test. The study found that during summer months (when temperatures in Australia were on average 28^0C), test positivity and f-Hb were significantly lower than during the cooler months. This did not affect

the detection of significant neoplasia, but did affect the detection rates of LRA as the f-Hb associated with these is near the cut-off. A further finding was a significant relationship between previous participation in screening and lower f-Hb and positivity. If screening programme providers are to fully realise the effectiveness of population screening further analysis of environmental, health and social factors is required.

In conclusion, uptake was higher than with the current screening algorithm and returned to usual levels when participants in the NHS Boards were invited with FIT were switched back to the current gFOBT/FIT algorithm. Untestable FIT due to spoiled, unlabelled and undated collection devices were lower than with the gFOBT/FIT algorithm. Calls to the helpline were lower. The quantitative nature of the analyses allows consistency of testing and permits modification of the f-Hb cut-off used - perhaps to allow for colonoscopy capacity. The increased uptake, ease of use for participants and favourable clinical outcomes of this feasibility study all support the introduction of FFLT in the SBoSP.

There is now evidence that FIT perform better than gFOBT (143) and an ever-growing number of countries are employing FIT as the screening modality. The European Group on Tumour Markers (EGTM) presented an updated paper detailing the current position in screening for CRC (27). There are nine different organisations that recommend national screening. However amongst this group there are a wide variety of modalities suggested. Furthermore, the International Agency for Research into Cancer (IARC) has recently published new guidelines for quality assurance in CRC screening, and these recommend the use of a quantitative FIT as the screening test of choice (31).

CHAPTER 5: f-Hb Partitioned by Sex and Age

5.1 Introduction

Screening the asymptomatic population for significant colorectal neoplasia, that is, CRC and HRA, has generated much interest over recent years. There are differences in approaches between countries (29), but the use of faecal tests is widespread, based on the findings of population-based RCT using gFOBT which demonstrated reductions in disease-specific mortality (43).

It has been documented in Scottish (46) (145), English (107) (153) and French (154) screening programmes that sex and age affect key performance indicators, including positivity. Although significant neoplasia is more common in men and older people, analysis of quantitative FIT data could be used to determine how f-Hb changes when certain important population characteristics are studied.

5.2 Aim of the Study

Current evidence supports the view that quantification of f-Hb in individuals is warranted (155). Laboratories commonly use sex and age partitioned reference values which describe 95% of the selected population, facilitating clinical interpretation of results. Guidance on how to do this is set out by the Clinical Laboratory Standards Institute (CLSI) in guidelines on how to define and determine reference intervals (156) (157).

Similarly, it may be that different cut-off or decision making tools using f-Hb could be used as the criteria for further investigation for different groups. In this study f-Hb, in a large cohort of individuals aged 50-74 years, the age group invited to take part in the SBoSP, were investigated. Potential

reference values were devised using the most recent international CLSI guidelines. The likely implications for CRC screening programmes of using decision limits based on different f-Hb in men and women and in different age quintiles were examined.

Estimation of f-Hb in a large group of ostensibly asymptomatic people prompted analysis of the data to see if patterns in sex, age and uptake of screening were present. The aim of this analysis was to investigate potential reference values. Those who returned a testable faecal sample were defined as the reference population. The usual SBoSP exclusion criteria were applied.

5.3 Methods

In order to determine whether the new FIT method would benefit a screening population a study of feasibility was undertaken in a selected group already eligible for the current SBoSP. Scotland has been offering screening to men and women between the ages 50 – 74 years since June 2007. The rollout proceeded as NHS Boards were able to meet the additional colonoscopy demand. The algorithm used consists of inviting participants with a gFOBT kit when potential participants reach 50 years of age or are in the target age range and become eligible on receiving a Community Health Index (CHI) number (Appendix 5). When the gFOBT kit is returned to the SBoSL and the test result is returned as positive the participant is offered colonoscopy. Where a weak positive or spoiled result is obtained the participant is sent a qualitative FIT test. If this is positive, the participant is referred for colonoscopy (Appendix 6). Where there is a negative test result, the participant is informed by letter and re-invited in two years.

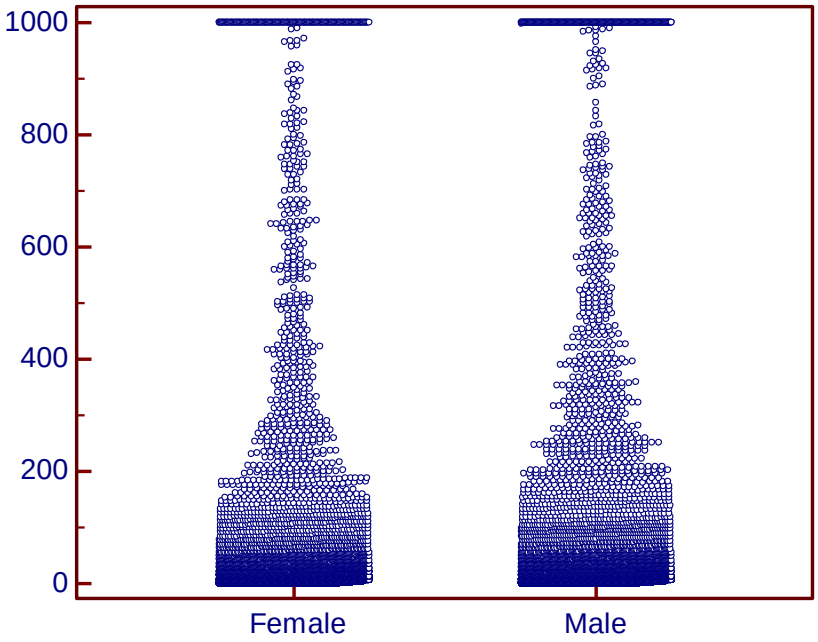
5.3.1 Statistical Analysis

Potential reference intervals were calculated for men and women and for both sexes in quintiles according to the CLSI Approved Guideline C28-A3c (157). The percentage of positive test results that would be obtained at a number of f-Hb were calculated from the cumulative percentage distributions. The number and percentages of individuals who would fall into the risk categories of Chen et al (158), namely, low, intermediate, high, and extremely high risk groups were defined using faecal concentrations of 1–19, 20–39, 40–79 and 80–99 ng Hb/ml buffer, respectively, were estimated from the distributions. MedCalc (MedCalc Software, Mariakerke Belgium) statistical software was used for all calculations.

5.4 Results

The number of people invited to participate in the study was 66,225: 48.8% were men and 51.2% women. Some participants returned a sample that was untestable due to being received more than 10 days after collection of faeces or spoiled in some way. Participants who returned such a sample were offered the chance to participate again with a further kit sent to them. The total number of specimen collection kits sent out during the study period was 68,041; this includes requests for repeat samples. The number who responded by returning a kit was 40,125 and thus the uptake was 60.6%. Of these, the reference sample group comprised 38,720 in total: 18058 men (46.6%) and 20,662 women (53.4%) who returned a testable sample and therefore received a negative or positive test result. Final uptake for the study group was therefore 58.5%. Figure 26 shows the distribution of Hb for men and women.

Figure 26 Distribution of f-Hb in 40,000 screening participants



It is usual practice in CRC screening to divide the population into quintiles. For men and women and for each of the age quintiles, none of the distributions of data were Gaussian (D'Agostino-Pearson test, $P < 0.0001$) and, in all cases, the coefficients of skewness and kurtosis were significantly greater than 1.00 ($P < 0.0001$). Tables 23 and 24 show that more females returned a kit than males (53.4%:46.6%). They also tabulate the distributions of f-Hb for men and women in five age groups with the conventional 95% CI. The 97.5% percentile represents the potential non-parametric upper reference limit of the 0.95 inter-fractile reference interval and 90% CI are given as recommended in Clinical and Laboratory Standards Institute (CLSI).

Table 23 Percentiles (with 95% CI) of f-Hb (ng Hb/ ml buffer) in men and potential upper reference limits (with 90% CI)

Age range (years)	N (%)	25.0%	50.0%	75.0%	90.0%	95.0%	97.5% URL
50-75	18058	0 (0-0)	1 (1-1)	12 (12-13)	67 (61-73)	184 (175-213)	**519 (468-575)**
50-54	4075 (21.9)	0 (0-0)	0 (0-0)	7 (7-8)	35 (29-41)	104 (91-127)	**281 (234-368)**
55-59	4160 (23.0)	0 (0-0)	0 (0-0)	9 (9-10)	48 (41-57)	154 (127-18t)	**415 (340-478)**
60-64	3489 (19.3)	0 (0-0)	1 (0-1)	12 (11-13)	62 (50-73)	185 (159-226)	**520 (411-663)**
65-69	3497 (19.4)	0 (0-0)	2 (2-3)	17 (15-20)	98 (87-112)	253 (298-319)	**713 (526-999)**
70-75	2837 (15.7)	0 (0-0)	3 (3-4)	21 (19-24)	125 (104-149)	347 (632-979)	**737 (647-927)**

Table 24 Percentiles (95% CI) of f-Hb (ng Hb/ ml buffer) in women and potential upper reference limits (with 90% CI)

Age range (years)	N (%)	25.0%	50.0%	75.0%	90.0%	95.0%	97.5% URL
50-75	20662	0 (0-0)	0 (0-0)	9 (8-9)	38 (35-41)	114 (106-121)	**283 (257-316)**
50-54	4543 (22.0)	0 (0-0)	0 (0-0)	6 (6-7)	23 (20-26)	68 (53-80)	**170 (142-248)**
55-59	4730 (22.9)	0 (0-0)	0 (0-0)	7 (6-8)	31 (27-34)	92 (76-115)	**244 (191-310)**
60-64	4058 (19.6)	0 (0-0)	0 (0-0)	8 (7-9)	34 (30-38)	101 (87-117)	**235 (184-313)**
65-69	3985 (19.3)	0 (0-0)	0 (0-0)	10 (10-11)	45 (40-53)	129 (112-151)	**317 (254-418)**
70-75	3346 (16.2)	0 (0-0)	0 (0-0)	15 (14-17)	80 (67-95)	190 (157-243)	**533 (409-675)**

Table 25 shows the positivity for men and women in the age quintiles at commonly used f-Hb cut-off (ng Hb/ml buffer) and the f-Hb cut-off to attain 2.0% positivity. More males than females were reported as having a positive test result using a 400 ng Hb/ml buffer cut-off (56.4%:43.6%).

Table 25 Positivity (%) at commonly used f-Hb (ng Hb/ml buffer) and cut-off to attain 2.0% positivity for men and women

Sex	Age range (years)	50	75	100	200	400	Cut-off to give 2.0% positivity
Men	50-75	11.6	9.3	7.8	4.8	3.0	680
	50-54	7.9	6.4	5,2	3.2	1.9	390
	55-59	9.7	6.7	6,6	4.0	2.6	525
	60-64	11.0	8.7	7.6	4.7	3.0	670
	65-69	15.0	11.8	9.7	6.0	4.5	950
	70-75	16.4	13.8	11.3	9.3	4.5	950
Women	50-75	8.4	6.7	5.5	3.1	2.0	400
	50-54	5.9	4.7	3.9	2.7	1.4	270
	55-59	7.2	5.7	4.8	2.8	1.9	350
	60-64	7.8	6.3	5.0	2.7	1.8	340
	65-69	9.4	7.4	6.3	3.3	2.2	430
	70-75	13.0	10.4	8.6	4.9	3.0	680

The number of individuals who would fall into the risk categories of Chen et al (158) namely, low, intermediate, high, and extremely high risk groups defined using faecal concentrations of 0–19, 20–39, 40–79 and 80–99 ng Hb/ml buffer, respectively are shown in Table 26.

Table 26 Percentage of individuals in low, intermediate, high, and extremely high risk groups (ng Hb/ml buffer)

Sex	Age range (years)	Low risk 0-19	Intermediate risk 20–39	High risk 40–79	Very high risk 80–99
Men	50-75	80.3	6.3	4.3	1.2
	50-54	85.6	5.1	3.1	1.0
	55-59	83.5	5.4	3.6	0.9
	60-64	80.2	6.9	4.4	0.9
	65-69	76.1	6.8	5.6	1.6
	70-75	73.9	7.9	5.4	3.1
Women	50-75	84.6	6.9	3.3	0.8
	50-54	88.9	4.3	2.4	0.5
	55-59	86.2	5.2	3.1	0.6
	60-64	85.6	5.6	2.8	1.0
	65-69	82.6	6.4	3.9	0.5
	70-75	77.9	7.1	4.9	1.4

5.5 Discussion and Conclusions

This study examined the f-Hb in a large group of ostensibly asymptomatic people aged 50-75 years. Table 24 shows that men have higher f-Hb than women in all age quintiles. This is already known, but the reason for this sex difference is not known. Brenner et al (159) have documented some possibilities. Men have higher blood Hb than women but the population of women here is likely to be mainly post-menopausal when such sex differences are not so apparent; in addition, use of aspirin and non-steroidal anti-inflammatory drugs might be more common in men than women. A plausible explanation is that colonic transit time is faster in men than in women. As a result, more degradation of f-Hb passed into the gut before passing of faeces takes place may occur in women since f-Hb is very unstable (14). In addition, f-Hb increases with age in both men and women but the reasons for this are unclear.

The question arises whether this large database should be used to create sex and age partitioned reference values for f-Hb. The non-parametric upper reference limits of the 0.95 inter-fractile reference interval were determined and, with 90% CI, are given in Table 22 and 23. The partitioned data was not able to be assessed against colonoscopy outcome in this study as it was not practicable within the constraints of the resources available.

In this study, the reference sample group is large and represents men and women generally regarded as asymptomatic. Although f-Hb is related to stage of neoplastic disease, (140) (160) (161) a number of individuals with high f-Hb had no abnormalities on colonoscopy and no evidence of disease. However, it is known that adenoma and early CRC are associated with few if any symptoms. Moreover, the fact that FIT are of considerable value in screening and that much early neoplastic disease is detected in individuals who participate in screening programmes makes it likely that our reference

sample group does contain a mix of healthy individuals, some taking drugs which stimulate bleeding into the gut or inhibit clotting and those with neoplastic disease and a number with other bowel conditions such as non-neoplastic polyps, inflammatory bowel disease, diverticular disease and haemorrhoids, all of which may cause bleeding into the gut. So, the question is – can population based reference values be set for f-Hb?

Inspection of the distributions of data shown in Table 22 and 23 suggest that, for both men and women, at least 90% have f-Hb less than 100 ng Hb/ml buffer, the manufacturer's recommended and most usual cut-off concentration used to refer individuals for the further investigation of bowel visualisation, usually colonoscopy. Indeed, more than half of the overall population screened have no detectable f-Hb. Thus, in spite of the accepted view that everyone has some blood in their faeces (162). It is postulated that the healthy individual has no detectable f-Hb and that any Hb present is not usual. In consequence, it is considered that population-based reference intervals for f-Hb, however established, would be of very limited value and that numerical decision limits should be applied in the screening setting and the data available here will be invaluable in setting these.

The positivity that would be found for men and women in the age quintiles at f-Hb cut-off commonly used in screening programmes are shown in Table 24. As expected, at any one concentration, the positivity found was higher for men than for women and was higher in older people than in younger people. This data can be used to predict positivity, the data given in Table 20 shows the f-Hb projected to gain a positivity of 2.1%, the ideal for the SBoSP since colonoscopy is a scarce resource in Scotland. Thus, although it is well known that men have more colorectal disease than women and such disease is much more common in the elderly, the use of a single f-Hb cut-off, as is current practice, in all screening programmes, is far from optimal. At any one cut-off, the distributions of f-Hb mean that more men

and older people will be referred for further investigation. The evidence presented here supports the recently stated view that there is a need for more tailored screening strategies (163).

Diagnostic laboratories commonly use sex and age partitioned reference values for other investigations. Similarly, it might be that different f-Hb cut-off should be used as the criteria for further investigation for men and women and different age quintiles. Because these data were generated in an ongoing national screening programme, colonoscopy was not performed in our entire reference sample group. However, what is needed to take this proposal forward is detailed evidence on the burden of disease in the groups labelled as positive and negative so as to examine objectively the use of sex partitioned decision limits in bowel screening programmes based on f-Hb.

An important study by Chen et al (158) investigated f-Hb below the usual cut-off concentration of 100 ng Hb/ml buffer as a predictor of incident colorectal neoplasia. f-Hb at first screening did predict subsequent risk of incident colorectal neoplasia. Table 25 shows the percentage of men and women in the age quintiles the population according to the low, intermediate, high and extremely high risk groups. As expected from the distributions of f-Hb, the percentage of men in each of the risk groups was higher than those of women and both sexes moved with increasing age towards a higher risk. These data could be used with advantage in screening programme planning: since the majority of individuals are in the low risk group, if screening was offered every three or four years for instance rather than every two years as is currently commonly done, resources would be saved. Even if, in addition, the high risk and very high risk individuals were invited annually rather than every two years, the overall number of positive screening test results, and thereby colonoscopies, would be lower than the usual current approach.

Risk scoring strategies exist for asymptomatic individuals which include sex and age (164), and for the symptomatic, risk scoring may include signs and symptoms, laboratory data, weight loss, diabetes and obesity (165). Now that automated quantitative FIT analyses are available, the evidence supports the concept that f-Hb could, be included in such scoring systems. This should prove a fruitful field of research for the future and is explored further in the conclusion. The statement made recently by Fraser (155) that, ‘with modern information technology, flair and imagination, risk adapted strategies could be adopted in future bowel screening programmes’ seems appropriate. The data presented here on associations between f-Hb and demographic characteristics should assist those involved in programme design.

CHAPTER 6: Using FIT in the Assessment of the Symptomatic Population

6.1 Introduction

Deciding which patients seen in primary care with symptoms suggestive of lower abdominal disease who will benefit from referral for investigation is difficult. As discussed recently by Manz et al (166), this is partly because the symptoms reported for colorectal diseases overlap considerably. Moreover, of such referred patients, only 22% to 37% actually have an important colorectal disease (23). There is growing evidence that faecal calprotectin is potentially useful for stratification of patients with lower abdominal symptoms, particularly as a rule-out test for inflammatory bowel disease (IBD), and an undetectable calprotectin concentration in a low-risk patient supports discharge of the patient without further investigation (76). There is little evidence to support guaiac-based faecal occult blood testing (gFOBT) in this decision-making context. Indeed, certain authoritative guidelines did state that these tests have no place in investigation of symptoms (19) (20). For these and other reasons, laboratories in the UK are very actively eliminating gFOBT from their repertoire and their use has been discouraged in other clinical settings (167).

In contrast gFOBT, has been demonstrated to reduce CRC mortality in randomised controlled trials of asymptomatic population screening (25). These results have been mirrored in bowel screening programmes (47) (66) that have been established following successful pilot studies. However, although gFOBT has some merit for use in structured population screening programmes, it also has major disadvantages (100) and many consider that gFOBT is now obsolete for use even in screening (168).

Newer tests for Hb detection have many advantages, are recommended for screening in current guidelines (31) and are being widely adopted. Quantitative FIT allow the measurement of f-Hb and have many benefits over qualitative FIT including the ability to vary the f-Hb cut-off used to refer for further investigation. As recently summarised by Rabeneck et al (120), there are many studies that document clinical outcomes in screening for colorectal neoplasia, which show that quantitative FIT are superior to gFOBT, particularly for adenoma detection. With regard to the use of these newer FIT in the assessment of patients with symptoms, a recent systematic review and meta-analysis on the value of the symptoms and additional diagnostic tests for CRC in primary care stated that, although results for qualitative FIT showed good diagnostic performance for CRC screening, evidence from primary care was lacking. It was suggested that high quality studies on their role in the diagnostic investigation in primary care are urgently needed (16).

6.2 Aim of the Study

The aim in this study was to examine the utility of f-Hb measurements, determined by a quantitative FIT, in deciding who of those that present in primary care with symptoms of lower abdominal disease would benefit from referral for endoscopy.

6.3 Methods

The performance characteristics of the OC-Sensor analyser are outlined in Chapter 3, briefly, the analysers were calibrated once per month, or when the latex LOT changed, with the calibrators provided. Each analytical run was preceded by analysis of two quality control materials at different Hb concentrations. The target values for the lots of materials used were set a

priori by 20 replicate analyses and a 1_2s rule used for acceptance or rejection of analytical runs.

This study was performed according to the STARD (Standards for the Reporting of Diagnostic accuracy studies) checklist as far as possible (138). This set of 25 reportable items is designed to improve quality and comparability across studies and is particularly useful for comparing analytical methods.

6.3.1. Evaluation Period

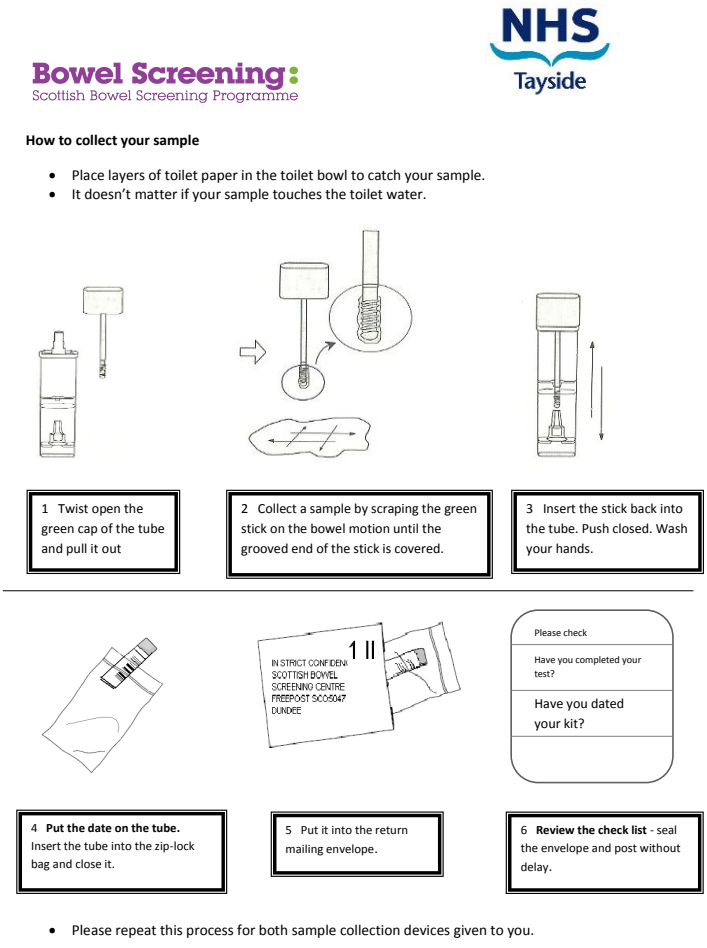
Patients who had been referred from primary care for endoscopic examination of the lower gastrointestinal tract in NHS Tayside from 01 February, 2010, to March 31, 2012, were invited to participate in this study. Reasons for referral were symptoms including rectal bleeding, change in bowel habit, IDA, abdominal pain, bloating, polyp/colorectal CRC surveillance, family history and assessment of inflammatory bowel disease (IBD). Patients who were under 16 years of age, unable to understand instructions or unable to consent were excluded. Those referred for investigation of a positive test result from the SBoSP were also excluded.

6.3.2. Participants and Sample Collection

Participants were recruited in the order that they appeared on the endoscopy appointment list, no intervention was made based on the f-Hb. A short letter of invitation to the study was included with their appointment letter (Appendix 7), along with a concise description of the rationale and aims of the study (Appendix 8). A research nurse followed up the invitation with a telephone call. If a patient was willing to take part, a sample collection device, written and pictorial instructions for collection of a faecal sample,

see Figure 27, a small labelled zip-lock plastic bag with integral absorbent material and a foil return mailing pouch were then mailed to the participant. The single faecal sample collection device is designed to collect ca. 10 mg freshly passed faeces using a serrated probe integral to the device cap into 2.0 ml of buffer in a tube. The label documented the participants unique study number numerically and in a barcode. The participant wrote the date of sample collection on the sample collection device and the information provided emphasised the need to post the sample back to the SBoSL immediately. The foil mailing pouches, with completed sample collection devices, were returned through the normal UK Post Office mail system by first class freepost.

Figure 27 Instructions for Use in the FIT in the symptomatic study



The single faecal sample collection device was distributed, and the label documented the participant's unique study number numerically and in a barcode format. The participant wrote the date of sample collection on the specimen collection device and the information provided emphasised the need to post the device back to the SBoSL immediately. The foil mailing pouches, with completed sample collection devices, were returned through the normal UK Post Office mail system by first class freepost.

6.3.3. Sample Handling

On return to the SBoSL the foil mailing pouches were opened, the barcode detailing the study number was used for analysis. Samples that were greater than 10 days from date of sample collection were termed expired and not tested further.

6.3.4. Measurement of f-Hb

The returned samples were analysed for f-Hb using one of two OC-Sensor Diana automated immunoturbidimetric analysers. Analyses were carried out in the SBoSL by trained staff whose major function is to perform faecal test analyses; the SBoSL has a comprehensive total quality management system and is accredited to ISO 15189 based standards by Clinical Pathology Accreditation (UK) Ltd. The analytical strategy and performance achieved have been detailed in chapter 3. The analysers were calibrated once per month with the calibrators provided. Each analytical run was preceded by analysis of two quality control materials at different Hb concentrations. The target values for the lots of materials used were set a priori by 20 replicate analyses and a 1_2s rule used for acceptance or

rejection of analytical runs. The combined imprecision obtained for the two analysers over the course of the study was CV < 3.8 %.

6.3.5. Reference Standard

Colonoscopy and FS were performed in the Endoscopy Unit, Ninewells Hospital and Medical School, Dundee, Scotland. Located within NHS Tayside, the Unit performs ca 1,000 procedures a year, and participates in the accreditation scheme of the Joint Accreditation Group on GI Endoscopy (JAG).

Data for endoscopy outcome and any subsequent pathology were downloaded from the appropriate NHS Tayside clinical IT systems. Data were collected on the quality of the investigation (quality of preparation, completeness) and on the results including number, size and localization of CRC, high-risk adenoma (HRA) and low-risk adenoma (LRA), and whether biopsy was performed. Full pathological data were collected on all excised and biopsy specimens including polyp type, presence or absence of malignancy, stage of any CRC and the characteristics of all adenomas. Following this, f-Hb from participants were collated into clinical outcome groups according to their most serious diagnosis. Assignment to the HRA group was dependent on the participant having three or more adenomas, or any adenoma with a maximum diameter of 1 cm or greater, as recommended by the BSG (as modified for general use in Scotland) (48). Finally, the reasons for referral for endoscopy were documented. The median f-Hb in all those who had a particular reason for referral were compared with the remainder of the patients in the overall study group.

6.3.6. Statistical Analysis

MedCalc (MedCalc Software, Mariakerke, Belgium) statistical software was used for all calculations. Data analyses after inclusion of 280 evaluable patients providing a maximum margin of error of 5% at 90% probability. Age and sex in the different groups were compared using the chi-squared test. PPV, NPV, sensitivity and specificity were calculated with 95% CI.

CRC, adenoma, particularly HRA, and IBD were defined as those conditions for which endoscopy was of particular benefit. Hyperplastic polyps (HPP), diverticular disease (DD), haemorrhoids and prolapse were considered as those conditions for which endoscopy was not of major clinical value, along with the finding of a normal large bowel. Analyses of the diagnostic accuracy of f-Hb were performed with participants grouped in four different combinations of these clinical outcomes as positive: Group 1 - all neoplasia plus IBD (CRC + HRA + LRA + IBD), Group 2 - significant neoplasia plus IBD (CRC + HRA + IBD), Group 3 - all neoplasia (CRC + HRA + LRA) and Group 4 – significant neoplasia (CRC + HRA). The remaining, with all other clinical endpoints and normal endoscopy (Group 5), were considered negative. f-Hb in various classes of symptoms were compared with the chi squared test. Probability of $p<0.05$ was considered significant.

Receiver operator characteristics (ROC) were used to determine the area under the curve (AUC) and analyse suitability as a diagnostic test (139).

6.4 Results

Ethical approval was granted by Tayside Research Ethics Committee (09/S1401/52, 04 June, 2009) and informed consent to participate was implied if a sample for FIT analysis was provided. In addition, the work was approved by the SBoSP Board and had Caldicott Guardian approval from NHS Tayside. 739 participants were invited to participate in the study, with 474 accepting and 265 declining. Although 291 completed both the FIT and

endoscopy, 11 of these were excluded either due to the FIT sample being completed after endoscopy or the sample not being dated and having arrived in the laboratory later than the endoscopy appointment date. Figure 28 shows the study flow with the number of patients invited, accepting, declining, undertaking both investigations, endoscopy, FIT, neither investigation and excluded.

Figure 28 Flow diagram showing number of participants completing FIT and colonoscopy following study invitation.

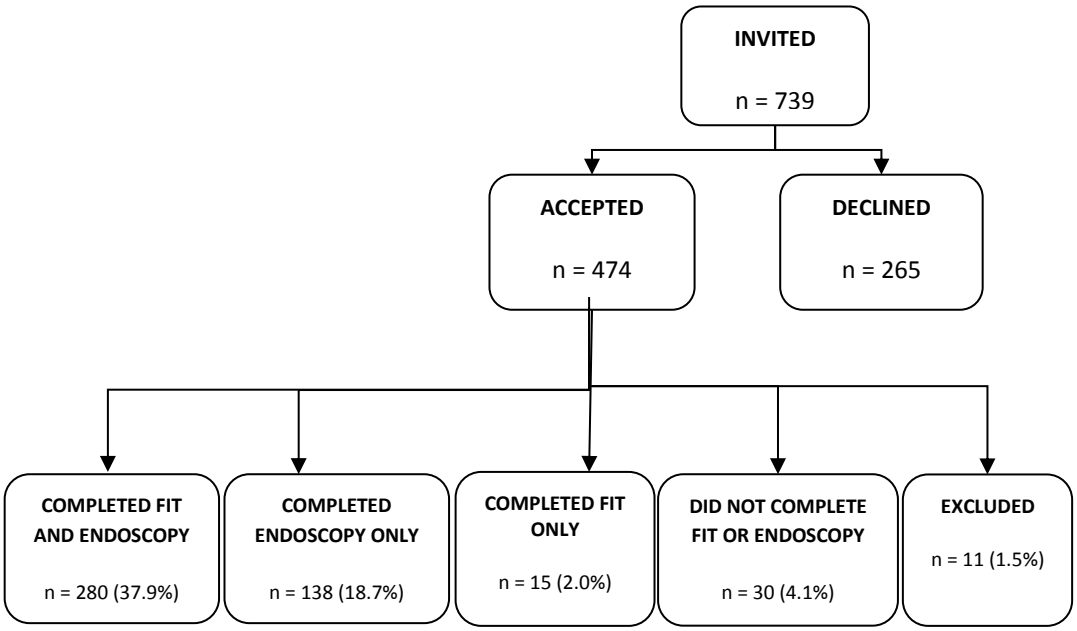


Table 27 shows the breakdown, of participants invited, accepting, declining and completing both, one of, or neither of the investigations, by sex and with median age compared with those excluded.

Table 27 Breakdown of number of participants completing FIT and colonoscopy following study invitation by sex and median age

	n	% of n	% of Invited	Median age (years)
Invited	739	100	100	59
Male	331	44.8	100	60
Female	408	55.2	100	59
Declined	265	100	35.9	57
Male	120	45.3	36.3	59
Female	145	54.7	35.5	56
Accepted	474	100	64.1	61
Male	211	44.5	63.7	60
Female	263	55.5	64.5	61
Completed neither	30	100	4.1	58
Male	15	50.0	4.5	56
Female	15	50.0	3.7	58
Completed FIT only	15	100	2.0	70
Male	7	46.7	2.1	57
Female	8	53.3	2.0	70
Completed endoscopy	138	100	18.7	57
Male	71	51.4	21.5	55
Female	67	48.6	16.4	58
Completed both	280	100	37.9	63
Male	113	40.4	34.1	64
Female	167	59.6	40.9	63

There were no statistically significant differences between the proportions of men and women when comparing those accepting and declining invitation. The ages of those invited ranged from 16 to 89 years. Comparison of the median age between groups showed that those who declined invitation were significantly younger than those who accepted ($p = 0.008$), and those who accepted and went on to complete both tests ($p < 0.0001$). In addition, those who underwent endoscopy only were younger than those who underwent both investigations ($p < 0.0001$). All of those who accepted the invitation but did not go on to complete the FIT were younger than all of those completing a FIT ($p < 0.0001$).

The number of days between the participant completing the FIT and date of endoscopy was investigated due to a possibility of change in disease status over time. Only two of the 280 patients included in those completing both investigations did not date their sample collection device, but the samples were received by the SBoSL four days and nine days prior to their endoscopy appointment date, and these were therefore included. The remaining 278 participants completed and dated their sample device prior to undergoing endoscopy within a range of 1 - 112 days: the median was nine days and inter-quartile range 4 - 20 days. The results of endoscopy for the 280 participants in the group who completed both investigations are documented in Table 28. 250 underwent colonoscopy and 30 flexible sigmoidoscopy. There were no adverse effects during this study.

Table 28 Clinical outcomes of participants completing both FIT and endoscopy

	n	%
CRC	6	2.1
All adenoma	54	19.3
High-risk adenoma (HRA)	23	8.2
Low-risk adenoma (LRA)	31	11.1
Hyperplastic polyps	12	4.3
Inflammatory bowel disease	26	9.3
Diverticular disease	68	24.3
Proctitis	2	0.7
Haemorrhoids	27	9.6
Prolapse	1	0.4
Normal	84	30.0
Total	280	100

To assess the relationship between f-Hb and severity of disease, ROC curve analysis was undertaken for those designated as positive in Groups 1 to 4 against all other data, defined as negative. Figures 29 - 32 show the ROC curves generated for each of the Groups.

The AUC for all Groups with the more important diseases classified as positive ranged from 0.734 to 0.671, suggesting only a “fair” to “poor” diagnostic test. The sensitivity and specificity were 62.1% or less and 88.1%

or less respectively at the optimum f-Hb cut-off. The overall optimum f-Hb for detection of significant disease was 51 ng Hb/ml buffer.

Figure 29 Receiver Operating Characteristic (ROC) curve with all neoplasia plus inflammatory bowel disease classed as disease

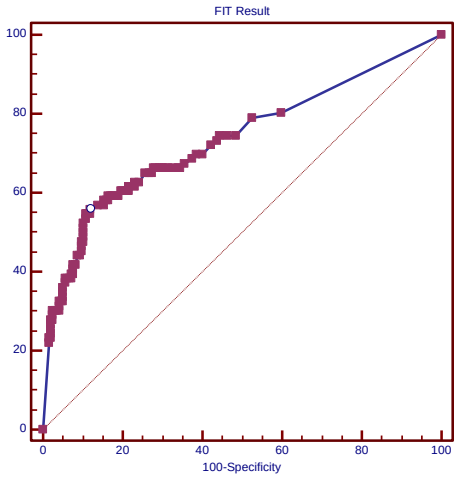


Figure 30 ROC curve with significant neoplasia plus inflammatory bowel disease classed as disease

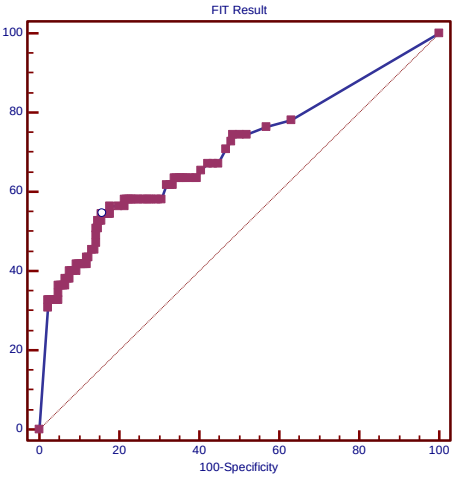


Figure 31 ROC curve with all neoplasia classed as disease

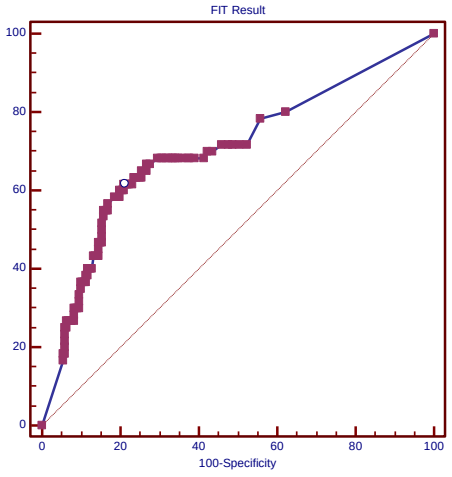
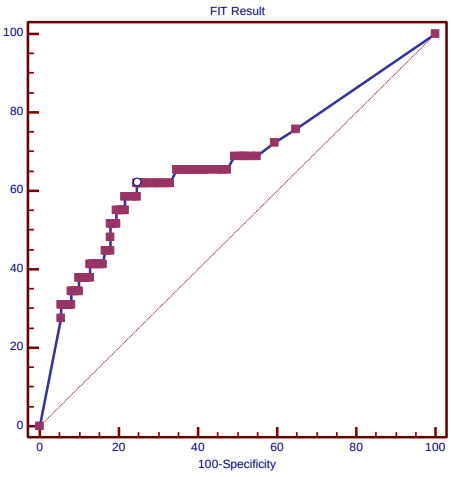


Figure 32 ROC curve with significant neoplasia classed as disease



The optimum f-Hb cut-off in the different clinical outcome groups are also shown together with 95% CI in Table 29. Each of the four groups described above to differently classify disease as positive (Groups 1 to 4) had

significantly higher optimum f-Hb cut-off than those in the Group (Group 5) classed as without significant disease, that is, classed as negative.

Table 29 Results of Receiver Operator Curve (ROC) analysis

Disease classification	Area under curve (95% CI)	Optimum f-Hb cut-off (ng Hb /ml buffer)	Percentage sensitivity (95% CI)	Percentage specificity (95% CI)
All neoplasia plus IBD	0.734 (0.678 - 0.785)	> 54	55.1 (44.7 – 66.5)	88.1 (82.7 – 92.3)
Significant neoplasia plus IBD	0.700 (0.643 – 0.753)	> 63	54.6 (40.6 – 68.0)	84.4 (79.0 – 88.9)
All neoplasia	0.702 (0.644 – 0.755)	> 41	61.7 (48.2 – 73.9)	79.1 (73.1 – 84.3)
Significant neoplasia	0.671 (0.612 – 0.725)	> 46	62.1 (42.3 – 79.3)	75.3 (69.5 – 80.5)
Overall		51		

Set out in Table 30 are the diagnostic characteristics of FIT at a f-Hb cut-off 50 ng Hb/ml buffer.

Table 30 Positive Predictive Value (PPV), Negative Predictive Value (NPV), sensitivity and specificity for different clinical outcome groups with a f-Hb > 50 ng Hb/ml buffer

Disease classification	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)
CRC	7.6	100.0	100.0	93.9
All adenoma	36.7	87.6	53.7	98.5
High-risk adenoma (HRA)	13.9	94.0	47.8	96.3
Low-risk adenoma (LRA)	22.8	93.5	58.1	97.7
Inflammatory Bowel Disease (IBD)	17.7	94.0	53.8	97.1
All neoplasia + IBD	62.0	81.6	57.0	99.0
Significant neoplasia + IBD	39.2	88.1	56.4	98.6
All neoplasia	44.3	87.6	58.3	98.7
Significant neoplasia	21.5	94.0	58.6	97.6
Other/no pathology	38.0	18.4	15.5	93.4

6.5 Discussion and Conclusions

This study is the first to evaluate the diagnostic accuracy of f-Hb measurements to discriminate between important colorectal disease and pathology of lesser clinical importance. Patients with CRC, HRA, LRA and IBD had significantly higher median f-Hb than those with findings considered to be of less clinical importance, namely HPP, DD, proctitis, haemorrhoids and prolapse. As expected, almost one-third of patients had no abnormality on endoscopy and these also had significantly lower f-Hb.

The decision to use one sample device in this study was based on literature related to screening populations. A recent evaluation of the clinical utility of using either one or two samples for the detection of disease in a symptomatic cohort reported that the diagnostic yield obtained by using the higher of two FIT samples (FIT/max) could also be obtained by using one FIT with a very low cut off (f-Hb ≥10 μg Hb/g faeces). This supports the strategy and findings presented in this work (169).

The data showed that all six patients with CRC had an elevated f-Hb with five having f-Hb greater than 1000 ng Hb/ ml buffer (the other having 346 ng Hb/ml buffer), but those with HRA, LRA and IBD had lower median f-Hb concentrations with wide confidence intervals, which is in accord with the concept that f-Hb is related to severity of disease (155). The highest AUC using ROC analysis was when CRC + HRA +LRA + IBD were considered as positive. In consequence, it might be considered that f-Hb measurement could be used as a rule-in test for important colorectal disease. However, the AUC for all Groups with the more important diseases classified as positive ranged from 0.734 to 0.671; traditionally, such AUC would suggest only a “fair” to “poor” diagnostic test.

Analysis of the ROC showed that the f-Hb cut-off that gave optimum sensitivity and specificity together was *circa* 50 ng Hb/ml buffer. At this concentration, sensitivity was 100% for CRC but only 47.8% to 58.6% for

other types of important colorectal disease. In consequence, f-Hb measurements may have limited value in the diagnosis of important colorectal disease in patients presenting in primary care with lower abdominal symptoms except for CRC. However, the NPV for CRC, HRA, LRA and IBD were 100.0%, 94.0%, 93.5% and 94.0% respectively. Thus, the important finding of this study was that f-Hb could be a very good "rule-out" test for important colorectal disease in those presenting with symptoms in primary care.

Endoscopy is a scarce resource and it is believed that f-Hb should be measured before referral from primary care: those with f-Hb ≥ 50 ng Hb/ml buffer could be referred for urgent investigation and those with f-Hb <50 ng Hb/ml buffer could be managed expectantly or referred for a specialist surgical or gastroenterology opinion. Since, in keeping with other studies (23) (166), 30% of those referred had no abnormality and also had a low f-Hb, this strategy would free up considerable endoscopy capacity for those who would benefit and potentially reduce waiting times. Patients often had more than one symptom reported. Symptoms or indications for endoscopy had no clear relationship with median f-Hb other than that performed for documented rectal bleeding and for family history. Thus, although f-Hb concentrations are associated with the severity of colorectal disease found, they do not correlate well with reported clinical symptoms.

Since the publication of the findings presented in this work, that low f-Hb could be used to rule out colonoscopy and help manage endoscopy resources, Rodriguez-Alonso et al. (170) reported on the effectiveness of CRC urgent referral guidelines (NICE and SIGN guidelines) compared to f-Hb in triaging symptomatic patients for colonoscopy. This was a prospective study of 1003 patients, 665 from primary care and 338 from secondary and tertiary care and colonoscopy was used as the gold standard. Analysis of the data showed that the number needed to scope to detect one case of cancer using NICE, SIGN, f-Hb ≥10 μg Hb/g faeces, and ≥15 μg Hb/g faeces (50 and 75 ng Hb/ml buffer) were 21, 24, 8 and 7 respectively and the

numbers needed to scope to detect one case of significant neoplasia were 6, 6, 3 and 3. Although this was a mixed primary, secondary and tertiary care cohort, these are encouraging results and add to the growing evidence base for using f-Hb in the triaging of symptomatic patients.

The major strength of the study presented in this Chapter is that the patients were recruited before endoscopy after being assessed as requiring this investigation. The setting of this study in primary care is important. Studies previously reported in secondary care do not deliver approaches and results that can be readily translated into a primary care setting. The approach adopted to testing in that setting needs to fulfill a number of requirements that have been addressed in this study. The primary care setting requires that diagnostic testing process must be easy for patients to comply with and provide a stable sample that can be transported easily to the laboratory for analysis. In addition the test should have good analytical practicability and reliability and provide results that are easy to interpret (171). The testing process for f-Hb used in the study presented here undoubtedly fulfil these requirements (100) (120) (167). The f-Hb test has an additional advantage of being viewed as inexpensive (approximately £3.00 for the sample collection device and analysis).

In Chapter 5 it is described that f-Hb is very dependent on sex and age in a screening population and suggested fixed decision-making limits be used based on sub sets of that group. On the basis of this study, two potential approaches to triaging patients in primary care could be developed. The use of centrally organised sample analysis or the use of simple POCT with an analytical detection limit stated to be 50 ng Hb/ml buffer (10 µg Hb/g faeces WEO recommended units) would appear to be viable approaches (172).

There is a growing interest in risk scores for use in detection of colorectal neoplasia in asymptomatic populations and there have been a number of approaches which have calculated independent risk factors that impact on the development of neoplasia (173) (177). It has been suggested that the

evidence supports the idea that f-Hb could, with advantage, be included in such risk scoring systems (22): this could prove a fruitful field of research in CRC screening for the future and is discussed further in the final chapter of this thesis. There are also scoring systems for use in assessment of those with symptoms: the diagnostic performance to identify symptomatic CRC compared with current referral guidance was recently investigated by Marshall et al (165) and Rodriguez-Alonso et al. (170). A more recent development has led to the creation of an Internet calculator, the QCancer® (Colorectal) risk calculator (www.qcancer.org/colorectal/) (174): the performance of this has been recently validated (175). In spite of the fact that little relationship was found between f-Hb and symptoms, it might be that such risk scoring systems for the symptomatic would be enhanced by addition of f-Hb, and that, even without further detailed clinical information, the evidence supports the view that knowledge of f-Hb would assist in the evaluation of symptomatic patients.

In conclusion, the study presented here indicates that f-Hb testing delivers a good rule-out test for clinically important colorectal disease. The outcome of the study would further suggest that application of f-Hb measurement to the symptomatic population in a primary care setting could enable delivery of a reduction in demand on endoscopy services.

CHAPTER 7: Highlights and Possible Future Work

7.1 Introduction

CRC is a significant cause of morbidity and mortality worldwide (2) (4). The presence of Hb in faeces can be used as a tumour marker, triggering further endoscopic investigation. A meta-analysis of four large RCT conducted in the 1990s was vital in providing evidence that a screening intervention, gFOBT, followed by colonoscopy for those who tested positive, significantly reduces disease specific mortality (43). However, these studies reported their findings decades ago and any test that performs better than gFOBT in terms of sensitivity and acceptability to the screened population should have a greater clinical impact than reported in these studies. Therefore the results achieved by the original RCT can be used to determine benchmarks whereby new screening modalities are investigated.

The primary aims of this work were:

- to determine if an automated quantitative FIT could improve outcomes for people who participate in the SBoSP
- to examine f-Hb data by sex and age and assess the implications and applications of any differences,
- to determine whether FIT can also be used to direct colonoscopy resources to those, referred from primary care, who will benefit most.

7.2 FIT as a First-Line Test

The effectiveness of a screening programme is influenced by a number of variables including the sensitivity of the test and it's acceptability to participants. In Chapter 4, it was reported that during the feasibility

evaluation, FFLT, only one faecal sample was collected rather than the six required with the gFOBT. A combination of this, the single sample collection, the simple sample collection device and new instructions for use may have been the main factors that contributed to an overall increase in uptake of the test in the study cohort. It was shown that uptake returned to previous levels when the gFOBT/FIT algorithm resumed in the evaluation NHs Boards. Analysis of uptake by deprivation quintile showed the greatest increase in uptake in the most deprived quintiles, and a significant increase in the most deprived men. SBoSP outcomes such as number of spoiled kits, calls to the helpline and turnaround time for results were all reduced. Adopting a 400 ng Hb/ml buffer (80 µg Hb/g faeces) cut-off to achieve a referral rate similar to that of SBoSP resulted in clinical outcomes with only subtle differences which may have been linked to screening round. Reducing the cut off to deliver increased test sensitivity would have the effect of increasing the number of colonoscopies performed within the screening programme.

The European Group on Tumour markers recommend the use of quantitative FIT with an adjustable cut-off for CRC screening. They identified that – 'iFOBT is superior to gFOBT with respect to detection rate and PPV for adenomas and CRC' (27). While there are many studies that document clinical outcomes that support the view that FIT are superior to gFOBT, non are in the context of national screening programmes. The studies used a low cut-off for FIT resulting in a higher positivity than could be resourced if adopted by the Scottish Bowels Screening programme (119) (143) (147). Findings presented in this thesis however demonstrate that use of FIT with a lower cut off could be adopted by the Scottish programme and deliver a number of benefits with a referral rate for colonoscopy comparable to that of the current gFOBT/FIT algorithm. Those benefits include improved outcomes in terms of screening uptake, compliance and turnaround time for results.

Further support for use of FIT as a frontline screening test comes from a prospective study undertaken in Saudi Arabia on 257 asymptomatic patients attending one of two tertiary hospitals (176). The authors reported on sensitivity and specificity of FOBT in screening for CRC. Each participant performed a once only gFOBT (Cologuard, Helena Laboratories, Beaumont, TX, USA) and qualitative FIT (RAPEPKT 313, DIAsource ImmunoAssays SA, Nivelles, Belgium 50 ng/ml buffer) and underwent a complete colonoscopy. The control group consisted of 20 cases of confirmed CRC who undertook the same regimen. What is interesting in this study is that patients were asymptomatic but everyone, with a positive and negative FOBT, was given a colonoscopy. In the asymptomatic cohort 37 participants were positive with one or both of the gFOBT and FIT, the remaining 194 were negative with both tests. The gFOBT and FIT positivity for the screening and control groups were 22.6%, 12.1%, 65% and 80% respectively. The sensitivity of gFOBT and FIT for the detection of CRC among the screening group was 50.0% (95% CI =6.76–93.24) and 75.0% (95% CI =19.41–99.37), and the specificity of gFOBT and FIT was 77.9% (95% CI =72.24–82.83) and 90.1% (95% CI =85.76–93.50).

This study adds to the growing evidence base indicating the effectiveness of FIT in screening for CRC. It is recognised that the Saudi Arabian study is relatively small in size with few cancers detected and wide confidence intervals bounding the reported sensitivities. Nevertheless the study lends some support to the case for using quantitative FIT in a screening programmes for CRC detection.

7.3 f-Hb is Related to Sex and Age

FIT measure f-Hb as continuous data variable and allow a cut-off to be chosen to give performance characteristics, such as screening test positivity, for objective assessment by screening programme organisers. In chapter 5, the investigation of the dispersions of f-Hb according to sex and

age was presented. It was shown, for the first time, that f-Hb increased with age in both men and women and that f-Hb was higher in men than in women in every age quintile. Differences amongst stratified data may prove useful in improving the efficiency of screening by enabling better targeted cut-off. Recent studies in Asia have suggested the possibilities of delivering individual-centred screening based on portioning of asymptomatic populations by age, gender, body mass index (BMI) (173) and the addition of further criteria such as, family history, smoking and presence or absence of diabetes mellitus (177). In the work undertaken here, the data gathered from all those who did attend colonoscopy demonstrated that, above the cut-off used increasing f-Hb is linked to increasing disease severity (178). It has now been shown, independently, that even below the conventional f-Hb cut-off can predict the risk of future significant neoplasia (158).

New data published on the use of different FIT analytical systems from various manufacturers confirms the findings presented in this thesis that f-Hb data can differ when partitioned according to age and gender (151). This holds true despite methods differing in terms of sample volumes, matrix effects and varying polyclonal antibody affinities for Hb and its degradation products. This is a consideration for programme organisers when procuring analytical equipment.

7.4 FIT in the Assessment of a Symptomatic Population

There is little evidence for the continued use of gFOBT in a spectrum of clinical settings. These tests have many serious disadvantages, should be considered obsolete and up until recently national guidelines have discouraged their use in primary care. In marked contrast, there is growing evidence that FIT would be useful, not only in screening for colorectal neoplasia, but also in the assessment of symptomatic patients.

The work described in Chapter 6 showed that f-Hb has very high NPV for significant colorectal disease (CRC, adenoma and IBD) in those patients who present in primary care with symptoms. FIT has the useful potential to be used as a rule-out test to help direct the available colonoscopy resource to those who would benefit most. In addition, data presented in this thesis showed all those patients with CRC had f-Hb above 50 ng Hb/ml faeces. A strategy of watch and wait for all those with f-Hb below this cut-off could potentially reduce endoscopy numbers by one third while ensuring that most disease would be picked up at colonoscopy.

One of the limitations of the work presented in this thesis is the small numbers. A larger, well organised, multicentre, prospective study of the diagnostic accuracy of f-Hb, NICE and SIGN guidelines for detecting CRC in symptomatic patients has been reported by Cubiella et al (179). This study used the Spanish COLONPREDICT cohort to determine the most reliable tool for predicting CRC and other significant colorectal diseases in a cohort referred for diagnostic colonoscopy. A total of 787 subjects met the selection criteria, 17.5% of patients were referred from primary care, 75.3% were from primary to secondary care and 7.1% were referred from secondary care. Assessment was based on f-Hb >100 ng Hb/ml buffer (>20 µg Hb/g faeces), NICE referral Criteria (20) and SIGN referral criteria (19). The diagnostic accuracy of the test for CRC, determined using AUC, showed FIT (0.88, 95% CI 0.84–0.92) was statistically higher than NICE (0.63, 95% CI 0.58–0.69) and SIGN (0.62, 95% CI 0.58–0.67) ($P < 0.001$) guidelines. In this study, use of FIT instead of NICE/SIGN referral criteria could reduce the number of colonoscopies required by 19.6% while increasing the number of CRC detected by 42.0%. This positive outcome supports the findings presented in this work.

Further support of the use of FIT as a tool for triage of patients presenting with symptoms comes from another Spanish study of patients, two thirds in

primary care and one third secondary and tertiary care, referred for colonoscopy (170). NICE and SIGN guidelines are compared to FIT with a cut-off set at 10 µg Hb/g faeces (50 ng Hb/ml buffer) and 15 µg Hb/g faeces (75 ng Hb/ml buffer). The number needed to scope was 21 and 24 respectively for the guidelines and 8 and 7 for the FIT. Using f-Hb above 10 µg Hb/g faeces as the criteria for referral to colonoscopy would only miss 0.1% of cancers and 5% of advanced adenomas. As suggested in this thesis, those people not referred could be placed under a system of 'watch and wait'.

In both the studies described above the use of SIGN and NICE guidelines improved the detection of significant neoplasia compared to no selection criteria, however the use of f-Hb further reduced the numbers needed to scope and increased the yield of significant neoplasia.

7.5 Adjusting the f-Hb Cut-Off within a Screening Programme

Three Scottish NHS Boards participated in the pilot rounds of screening (46) and from 2007 all NHS Boards are now participating in the SBoSP (66). In each pilot round the number of CRC detected and the PPV decreased. It is likely that this phenomenon is caused by screening out prevalent CRC and mainly finding incident CRC in the subsequent rounds and this still continues in the programme (108). If a screening strategy were to continue with the same f-Hb cut-off for many rounds, as the current programme does, the f-Hb within the screened population is likely to fall below the detection limit of the gFOBT and the number of CRC detected by this method will decrease. Introduction of a quantitative test where the end user sets the cut-off could resolve this issue as the f-Hb cut-off could be reduced to maintain a constant positivity which would allow those with less and less f-Hb to be referred for colonoscopy. This strategy could lead to a reduction in IC as

participants with slightly less f-Hb (women and younger participants) are sent for colonoscopy.

Examination of the three rounds of Scottish pilot data revealed the interval cancer rate increased with each round, from 31.2 to 58.9% (46). Data from Dutch studies show that using a low cut-off and referring high numbers of participants for colonoscopy increases the PPV for CRC (118). It was anticipated that a benefit of using a FIT strategy in the SBoSP would be a statistically significant increase in CRC detected; however, with the positivity set the same as the current gFOBT/FIT algorithm there was no significant increase in CRC detection. At this point it seems unlikely there will be a reduction in the number of IC in the evaluation group. Data regarding IC in the FFLT cohort will be very interesting when they are available.

One of the most translatable findings of this work is that f-Hb is higher in men then in women and increases with age. This presents an opportunity to review how the current SBoSP is delivered. It is known that incidence of CRC is higher in men, that f-Hb is higher in men and, within the context of the pilot rounds, IC were higher in women than men (33).

The quantitative nature of FIT allows for the f-Hb cut-off to be selected to give certain predesigned outcomes. If reduction of IC in women was an aim, the cut-off for women could be set lower than that used for men. If disease detection in younger ages was important the f-Hb cut-off for 50 – 60 years could be lowered or for over 75 years could be increased. Even though these changes could address inherent biases in the current dichotomous system, any changes would be dependent on not disadvantaging specific groups.

Currently, many countries (including Scotland) are planning on using a single cut-off for referral for colonoscopy with the inherent sex biases, deprivation gradient and high IC rates. There is technology available already

that could allocate each individual a cut-off based on age and gender which are contained within each CHI number so individualising CRC screening and allowing monitoring of f-Hb with time. It would be gratifying to see more imaginative approaches being taken in the near future.

7.6 Use of f-Hb in Risk Scoring

It has been documented in this work that sex, age and deprivation are significant predictors of disease, and as such use of quantitative FIT as a dichotomous test does not fully utilise the benefits of f-Hb. Given that many inter-relationships in the development of CRC exist, there are opportunities to develop multivariate risk tools for screening and triage of symptomatic patients. This would allow an individualised approach to healthcare, and targeting of those who require endoscopy and reduce the pressure on this resource.

A number of risk scoring models have been developed and published recently which demonstrate that stratifying risk, even without f-Hb, works well in asymptomatic populations. Since these studies are likely to inform programmatic effectiveness they are described in further detail here. Wells et al. (180) reported on the development of CRC-PRO in the USA. Case-control data was used to populate and internally validate a risk prediction model. Interestingly the variables used for men and women differed and increased BMI affected men and women differently. The impact of different variables were ranked in two tables relating to men and women; the most important variable in men and women was age which also had the strongest affect over any other variable. This model demonstrates the importance of using more than just one variable such as age as the spread of risk within one age group is the result of many factors e.g., a 50 year old woman's 10 year risk of CRC may be between 0.2% and 2.0% dependent on the other variables included in this model. The ability to predict risk of developing

CRC could enable decisions to be made as to whether an individual may benefit from early endoscopic intervention or more frequent screening or longer screening intervals are adopted. The model gave good prediction accuracy. However, one of the underlying principles of a risk prediction model is that it should be simple enough to be used in clinical care.

In 2015, Garcia et al. (180) published an analysis of f-Hb in relation to age, gender and clinical outcome in a Spanish screening cohort. Within Spain CRC screening is managed at a regional level and in 2010/2011 all regions operating a screening programme (12 out of 17 regions) switched to FIT, although they still report test results as positive or negative. This study reported on 1406 subjects who completed a colonoscopy after a positive (100 ng/ml buffer) screening test result. Similar to the findings presented in this thesis, increased f-Hb was associated with increased disease and f-Hb was significantly higher in men than in women. However, in this study increased age was not associated with increased f-Hb. Logistic regression, adjusted for age and sex, showed high-grade dysplasia, villous histology, distal location and increasing size were all associated with increasing f-Hb. All of this data was analysed with different cut-off to examine the effect of increasing the f-Hb cut-off on clinical outcomes. Increasing the f-Hb cut-off by 150 ng/ml buffer would reduce the positivity to 3%, but miss 15.9% CRC and 33.4% HRA. It was observed that as f-Hb cut-off is increased, PPV increased: however, there was more likelihood of missing lesions.

Chen et al. (182) developed a risk scoring system for asymptomatic average-risk populations with Han nationality in the Jiangsu Province. A total of 905 cases met the study criteria. In this study multivariate analysis showed age, gender, coronary heart disease, egg intake and defecation frequency were independent variables associated with increased risk of neoplasia. These were used to populate the scoring system: patients were scored from 0 – 10 with 2.5 being the cut-off for referral colonoscopy. Using

this scoring system reduced the number of colonoscopies required by 45.5% and missed three colorectal neoplasms, the NPV was 99.3%.

In Singapore, Yeoh et al. (164) reviewed data from asymptomatic screening colonoscopies across Asia. One third of the cohort were used to develop the model and the remainder to validate it. The Asia Pacific Colorectal Screening (APCS) score used age, gender, family history and history of smoking. The scoring system had a range of 0-7 points and risk was stratified as average (0-1), moderate (2-3) and high (4-7). Advanced neoplasm was used as the end point as polypectomy gives secondary benefits within screening. No advanced neoplasms were found in the average risk group, who comprised 19.2% of the development cohort. However, in the validation cohort, average risk individuals comprised 29.5% of the cohort and had seven advanced neoplasms (12%).

The APCS was tested with additional metabolic syndrome (MS) components (obesity, hyperlipidemia, hypertension and diabetes) by Wang (183). The additional MS components increased the amount of colorectal neoplasia that was detected, but this was not significant when compared to the APCS system.

These examples of risk scoring systems for use in asymptomatic populations are valuable because they offer the opportunity to reduce the number of colonoscopies without a significant loss in sensitivity. Many studies used the Hosmer-Lemeshow goodness of fit statistic to determine the reliability of the risk model in validation cohorts. If the result is <0.05 the model is not good and should be rejected, above 0.5 and it is a good fit. For the studies that completed this test the results were 0.49 (164), 0.174 (182). Wells was not able to test the CRC-PRO model with a different cohort since adequate data had not been collected. However, there are a number of limitations in that there are many variables that affect the risk of developing CRC and deciding which to include and how many variables to include will

affect the ability of the system to determine risk. In order to manage data some continuous variables must be portioned e.g. age and this may reduce accuracy and these models relied on an element of self-reporting which is not always accurate. In addition, models are not likely to be transferable across geography and populations since race and ethnicity affect CRC risk.

Addition of f-Hb in a scoring system should increase its ability to predict risk. This approach was taken by Rodreguiz-Alonso et al. (170) who used the data from a large cohort of symptom based referrals to construct a model to score risk of significant neoplasia and inform referral for colonoscopy. In this model age, gender and f-Hb were selected as they were linked to CRC in the analysis of the outcomes. The cohort were split into two groups, two-thirds of the group were used to develop the risk scoring system and the other third were used to internally validate the system. Using this algorithm only 36.4% of the presenting population would need to be referred, all cancers would be detected and only 5.0% of advanced adenomas would be undetected.

These studies demonstrate that reliable and accessible factors such as age, gender, BMI, smoking history and family history can be used to populate CRC risk models. However, the use of f-Hb should be included in any risk scoring model, in screening and symptomatic populations. In Scotland the screening programme has access to data on age and gender and so a risk scoring model could be developed with the existing data from the FFLT evaluation through the use of the CHI number and using established collaborative links developed with many with European countries, it would be possible to test it using similar screening data. Experience gained during the development of this work suggests that FIT is capable, within the context of risk scoring, of giving a result that has clinical characteristics similar to those seen in many diagnostic tests and evidence suggests that the number and frequency of colonoscopies could be reduced using this approach.

7.7 What Screening Interval should be used in a FIT Based Algorithm?

Patients who undergo colonoscopy and have significant neoplasia detected are placed under surveillance, meaning further colonoscopy at 1, 3 or 5 year intervals. Family history may place a person in a ‘high risk’ category where frequent colonoscopy is recommended. This approach is laid out in the BSG Guidelines (48) and is a balance of benefits and risks as colonoscopy is a limited resource and may have complications.

The FIT developments outlined in this work present an opportunity to rationalise and prioritise colonoscopy to those people who have evidence of need, decreasing waiting times and therefore improving outcomes. However, there is still very little research as to the interval that should be used between screening rounds. A recently published study investigated whether an offer of an annual, biennial or triannual single sample algorithm should be used and found that using an interval 1-3 years did not affect the detection of advanced neoplasia (124). Evidence is being gathered as to whether it is more beneficial to screen a small age group intensively with a low f-Hb cut-off or whether to screen for longer with a higher f-Hb cut-off. Early work indicates the benefits of an annual FIT strategy with a higher f-Hb cut-off when colonoscopy resources are limited (125). However, it has also been found that lowering the f-Hb cut-off improves detection of advanced adenomas rather than CRC (122).

Quantitative FIT results could be used to build a dataset to determine whether f-Hb at 50 years of age (index f-Hb) could be used indicate future risk of neoplasia, or a number of consecutive negative test results may

indicate the individual had less risk of developing significant neoplasia and that the screening/surveillance interval could be increased or if, following the index f-Hb there is an increase in f-Hb, a referral for bowel visualisation could be conducted.

7.8 Conclusions

There is significant potential for quantitative FIT to contribute to improved outcomes for people with undetected CRC in screening and symptomatic populations. The data presented in this work relate to the use of automated quantitative FIT in these two areas of interest; replacement of the gFOBT/FIT modality for early detection in the SBoSP and to triage those presenting in primary care.

A significant part of the work in this thesis relates to the progression from a binary system of reporting analytical outcomes to development of risk scores underpinned by quantitative FIT results and weighted risk factors such as age and gender. It is hoped that outputs from this work are translated into clinical practice and that, in the near future, patients in Scotland, in all relevant clinical settings, have access to investigations based on level of risk to provide high quality evidence based, healthcare.

References

1. International Agency for Research on Cancer. Colorectal Cancer Incidence, Mortality and Prevalence Worldwide in 2012 Factsheet. http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx?cancer=colorectal [Accessed: 05 May 2015]

2. Cancer Research UK, Bowel Cancer Incidence Statistics. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence [Accessed: 01 October 2015]

3. Young A, Hobbs R, Kerr D. ABC of Colorectal Cancer. Second Edition. Chapter 1 Pathways to Carcinogensis. Chichester. John Wiley & Sons. 2011

4. Cancer Research UK, Bowel Cancer Survival Statistics. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival [Accessed: 01 October 2015]

5. Libby G, Brewster DH, McClements PL, et al. The impact of population-based faecal occult blood test screening on colorectal cancer mortality: a matched cohort study. *British Journal of Cancer* 2012;107: 255-259.

6. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult blood screening on the incidence of colorectal cancer. *New England Journal of Medicine* 2000;343: 1603-1607.

7. Fearon ER and Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;61: 759-767.

8. Leslie A, Carey FA, Pratt NR, et al. The colorectal adenoma-carcinoma sequence. *British Journal of Surgery* 2002;89: 845-860.

9. Conlin A, Smith G, Carey FA, et al. Colorectal cancer: The prognostic significance of K-ras, p53, and APC mutations in colorectal carcinoma. *Gut* 2005;54: 1283-1286.

10. Stryker SJ, Wolff BG, Culp CE, et al. Natural history of untreated colonic polyps. *Gastroenterology* 1987;93: 1009-1013.

11. Bedenne L, Faivre J, Boutron C, et al. Adenoma-carcinoma sequence or “de novo” carcinogenesis? *Cancer* 1992;69: 883-888.

12. Stroehlein, JR, Fairbanks VF, McGill DB, et al. Hemoccult detection of fecal occult blood quantitated by radioassay. *The American Journal of Digestive Diseases* 1976;21: 841-844.

13. Kufe, DW. Pollock, RE. Weichselbaum, RR. et al. Holland-Frei Cancer Medicine. Hamilton: BC Decker. 2003.

14. Lampe JW, Fredstrom SB, Slavin JL, et al. Sex differences in colonic function: a randomized trial. *Gut* 1993;34: 531-536.

15. Rozen P, Young GP, Levin B, et al. Colorectal Cancer in Clinical Practice. London: Martin Dunitz. 2002.

16. Jellema P, van der Windt DAWM, Bruinvels DJ, et al. Value of symptoms and additional diagnostic tests for colorectal cancer in primary care: systematic review and meta-analysis. *British Medical Journal* 2010;340: c1269. doi 10.1136/bmj.c1269 [Accessed: 23 November 2014]

17. Panzuto F, Chiriatti A, Bevilacqua S, et al. Symptom-based approach to colorectal cancer: survey of primary care physicians in Italy. *Digestive and Liver Disease* 2003;35: 869-975

18. Goddard AF, James MW, McIntyre AS, et al. on behalf of the British Society of Gastroenterology. Guidelines for the management of iron deficiency anaemia. *Gut* 2011;60: 1309-1316.

19. Healthcare Improvement Scotland. Scottish Intercollegiate Guidelines Network. Diagnosis and Management of colorectal cancer. http://www.sign.ac.uk/pdf/sign126.pdf [Accessed: 05 May 2015]

20. NICE. National Institute for Health and Clinical Excellence. Colorectal cancer - the diagnosis and management of colorectal cancer - CG131. http://guidance.nice.org.uk/CG131 [Accessed: 05 May 2015]

21. Healthcare Improvement Scotland. SIGN. Scottish Intercollegiate Guidelines Network. Management of epithelial ovarian cancer. SIGN 135. http://www.sign.ac.uk/guidelines/fulltext/135/index.html [Accessed: 6 November 2015]

22. Fraser, CG. A future for faecal haemoglobin measurements in the medical laboratory. *Annals of Clinical Biochemistry* 2012;49: 518-526.

23. Kok L, Elias SJ, Witteman BJM, et al. Diagnostic accuracy of point-of-care fecal calprotectin and immunochemical occult blood tests for diagnosis of organic bowel disease in primary care: The cost-effectiveness of a decision rule for abdominal complaints in primary care (CEDAR) study. *Clinical Chemistry* 2012;58: 989-998.

24. Wilson JM and Jungner F. Principles and practice of screening for disease. World Health Organization, 1968. Public Health Papers No 34.

25. Hewitson P, Glasziou P, Watson E, et al. Cochrane systematic review of colorectal cancer screening using the faecal occult blood test (Hemoccult): an update. *American Journal of Gastroenterology.* 2008;103: 1541-1549.

26. Benson VS, Patnick J, Davies AK, et al. on behalf of the International Colorectal Cancer Screening Network.Colorectal Cancer Screening: A comparison of 35 intiatives in 17 countries. *International Journal of Cancer* 2008;122: 1357-1367.

27. Duffy MJ, van Rossum LGM, van Turenhout ST, et al. Use of faecal markers in screening for colorectal neoplasia: A European Group on Tumor Markers (EGTM) position paper. *International Journal of Cancer* 2011;128; 3-11.

28. Levin, B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: A joint guideline from the American Cancer Society, the US Multi-Society Taskforce on colorectal cancer and, the American College of Radiology. *CA A Cancer Journal for Clinicans* 2008;58: 130-160.

29. Hoff G and Dominitz JA. Contrasting US and European approaches to colorectal cancer screening: which is best? *Gut* 2010;59: 407-414.

30. Nadel MR, Shapiro JA, Klabunde CN, et al. A national survey of primary care physicians' methods for screening for fecal occult blood. *Annals of Internal Medicine* 2005; 142:86-94.

31. Patnick J, von Karsa L, Segnan N eds. European Guidelines for Quality Assurance in Colorectal Cancer Screening and Diagnosis. First Edition. Luxembourg. Publications Office of the European Union. 2010.

32. Health Services Management Centre of the University of Birmingham and Midlands and Lancashire Comissioning Support Unit. Scoping the future: evaluation of endoscopy capacity across the NHS in england. 2015. http://www.birmingham.ac.uk/schools/social-policy/departments/health-services-management-centre/research/projects/2015/evaluation-of-endoscopy-capacity-across-the-nhs-in-england.aspx [accessed 02 October 2015]

33. Steele RJC, McClements P, Watling C, et al. Interval cancers in a FOBT-based colorectal cancer screening programme: implications for stage; gender and tumour site. *Gut* 2012;61: 576-581.

34. Mandel JS, Bon, JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. *New England Journal of Medicine* 1993;328: 1365-1371.

35. Hardcastle JD, Chamberlain JO, Robinson MHE, et al. Randomised controlled trial of faecal occult blood screening for colorectal cancer. *Lancet* 1996;348: 1472-1477.

36. Scholefield JH, Moss SM, Mangham CM, et al. Nottingham trial of faecal occult blood testing for colorectal cancer; a 20 year follow up. *Gut* 2011;61: 1036-1040.

37. Robinson MHE, Thomas WM, Hardcastle JD, et al. Change towards earlier stage at presentation of colorectal cancer. *British Journal of Surgery* 1993;80: 1610-12.

38. Scholefield JH, Robinson MH, Mangham CM, et al. Screening for colorectal cancer reduces emergency admissions. *European Journal of Surgery and Oncology* 1998;24: 47-50.

39. Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal occult blood test. *Lancet* 1996;348: 1467-1471.

40. Jorgensen OD, Kronborg O and Fenger C. A randomised study of screening for colorectal cancer using feacal occult blood testing: results after 13 years and seven biennial screening rounds. *Gut* 2002;50: 29-32.

41. Faivre J, Dancourt V, Lejeune C, et al. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. *Gastroenterology* 2004;126: 1674-1680.

42. Lindholm E, Brevinge H and Haglind E. Survival benefit in a randomised clinical trial of faecal occult blood screening for colorectal cancer. *British Journal of Surgery* 2008;95: 1029-1036.

43. Towler B, Irwig L, Glasziou P, et al. A systematic review of the effects of screening for colorectal cancer using the faecal occult blood test, Hemoccult. *British Medical Journal* 1998;317: 559-565.

44. Steele RJC, Parker R, Patnick J, et al. A demonstration pilot for colorectal cancer screening in the United Kingdom: a new concept in the

introduction of health care strategies. *Journal of Medical Screening* 2001;8: 197-202.

45. UK Colorectal Cancer Screening Pilot Group. Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom. *British Medical Journal* 2004;329: 133-135.

46. Steele RJC, McClements PL, Libby G, et al. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. *Gut* 2009;58: 530-535.

47. Logan RFA, Patnick J, Nickerson C, et al. Outcomes of the bowel cancer screening programme (BCSP) in England after the first 1 million tests. *Gut* 2012;61: 1439-1446.

48. Cairns SR, Scholefield JH, Steele RJ, et al. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002) *Gut* 2010;59: 666-690.

49. Cancer Care Ontario. Colorectal Cancer. https://www.cancercare.on.ca/cms/one.aspx?portalId=1377&pageId=67772 [Accessed: 10 May 2014].

50. Holme Ø, Løberg M, Kalager M, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality. *The Journal of the American Medical Association* 2014; 312(6): 606-615.

51. Atkin WS,Edwards R, Kralj-Hans I, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised trial. *The Lancet* 2010;375: 1624-1633.

52. Segnan C, Ameroli P, Bonelli L, et al. Once-only sigmoidoscopy in colorectal cancer screening: follow up findings of the Italian randomised control trial - SCORE. 2011 *Journal of the National Cancer Institute* 2011;103: 1310-1322.

53. Weissfeld JL, Schoen RE, Pinsky PF, et al. Flexible sigmoidoscopy in the PLCO cancer screening trial: Results from the baseline screening examination of a randomized trial. *Journal of the National Cancer Institute* 2005;97: 985-997.

54. Schoen RE, Pinsky PF, Weissfeld JL, et al. for PLCO Group. Results of repeat sigmoidoscopy 3 years after negative sigmoidoscopy. *Journal of the American Medical Association* 2003;290: 41-48.

55. Nelson RL and Scwhartz A. A survey of individual preference for colorectal cancer screening technique. *BioMed Central Cancer* 2004;4: 76.

56. Kronborg O and Regula J. Population screening for colorectal cancer: Advantages and drawbacks. *Digestive Diseases* 2007;25: 270-273.

57. Winawer S, Zauber AG, MayNah H, et al. Prevention of colorectal cancer by colonoscopic polypectomy. *New England Journal of Medicine* 1993;329: 1977-1981.

58. Brenner H, Chang-Claude J, Seiler CM, et al. Interval cancers after negative colonoscopy: population-based case-control study. *Gut* 2011;61: 1576-1582.

59. van Gelder RE, Jung Nio C, Florie J, et al. Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer. *Gastroenterology* 2004;127: 41-48.

60. Stoop EM, De Hann M, de Wijkerslooth TR, et al. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. *Lancet Oncology* 2012;13: 55-64.

61. Pickhardt PJ. Strong evidence in support of CT colonography screening. *The Lancet Oncology* 2012;13: 6-7.

62. Senore C, Ederle A, Benazzato L, et al. Offering people a choice for colorectal cancer screening. *Gut* 2012;62: 735-740.

63. Libby G, Bray J, Champion J, et al. Pre-notification increases uptake of colorectal cancer screening in all demographic groups: a randomized controlled trial. *Journal of Medical Screening* 2011;18: 24–29.

64. Cole SR, Smith A, Wilson C, et al. An advance notification letter increases participation in colorectal cancer screening. *Journal of Medical Screening* 2007;14: 73-75.

65. Steele RJC, Kostorou I, McClements P, et al. Effect of repeated invitations on uptake of colorectal cancer screening using faecal occult blood testing: analysis of prevalence and incidence screening. *British Medical Journal* 2010;341;c5531.

66. Fraser CG, Digby J, McDonald PJ, et al. Experience with a two-tier reflex gFOBT/FIT strategy in a national bowel screening programme. *Journal of Medical Screening* 2012;19: 8-13.

67. Robinson MHE, Hardcastle JD, Moss SM, et al. The risks of screening: data from the Nottingham randomised controlled trial of faecal occult blood screening for colorectal cancer. *Gut* 1999;45: 588-592.

68. Duffy SW, Tabar L, Olsen AH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *Journal of Medical Screening* 2010;17: 25–30.

69. Diehl F, Schmidt K, Durkee KH, et al. Analysis of mutations in DNA isolated from plasma and stool of colorectal cancer catients. *Gastroenterology* 2008;135: 489-498.

70. Hundt S, Haug U and Brenner H. Blood markers for early detection of colorectal cancer: A systematic review. *Cancer Epidemiology, Biomarkers and Prevention* 2007;16: 1935–1953.

71. deVos T, Tetzner R, Weiss G, et al. Circulating methylated SEPT9 DNA in plasma is a biomarker for colorectal cancer. *Clinical Chemistry* 2009;55: 1337-1346

72. Church TR, Wandell M, Lofton-Day C, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer. *Gut* 2014;63: 317-325

73. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Fecal DNA versus Fecal Occult Blood for colorectal-cancer screening in an average-risk population. *New England Journal of Medicine* 2004;351: 2704-2714.

74. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. *The New England Journal of Medicine* 2014;370: 1287-1297

75. Osborn NK and Alquihist DA. Stool screening for colorectal cancer: Molecular approaches. *Gastroenterology* 2005;128: 192-206.

76. Ayling RM. New faecal tests in gastroenterology. *Annals of Clinical Biochemistry* 2012;49: 44-54.

77. Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer. *New England Journal of Medicine* 2013; 369; 1106-1114

78. Saito H. Screening for colorectal cancer by immunochemical fecal occult blood testing. *Japanese Journal of Cancer Research* 1996;87: 1011-1024.

79. Fujirebo. Magstream HT. http://www.fujirebio.co.jp/english/product/immunological [Accessed: 16 August 2013].

80. Nakama H, Kamijo N, Miyata K, et al. Sensitivity and specificty of several immunochemical tests for colorectal cancer. *Hepatogastroenterology* 1998;45: 1579-1582.

81. Ostrow, JD. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Chapter 98: Tests for Fecal Occult Blood. Boston. Hall Butterworths. 1990.

82. Immunostics Inc. Technical Info. Immunostics Inc. http://www.immunostics.com/#!technical/cipy [Accessed: 17 Nov 2013].

83. Greegor DH. Occult blood testing for detection of asymptomatic colon cancer. *Cancer* 1971;28: 131-134.

84. Halloran S, Seaman H and Young M. Expert Working Group: Standardisation: Summary tables all manufacturers. World Endoscoy Organisation, San Diego. 28 May 2012.

85. Cytodiagnostics. Lateral flow Assay. Cytodiagnostics. http://www.cytodiagnostics.com/lateral-flow-immunoassays.php [Accessed: 18 July 2014].

86. Saito H. Colorectal cancer screening using immunochemical faecal occult blood testing in Japan. *Journal of Medical Screening* 2006;13: S6-S7.

87. Beg M, Singh MK, Saraswat MK, et al. Occult gastro-intestinal bleeding: Detection, interpretation, and evaluation. *Journal, Indian Academy of Clinical Medicine* 2002;3: 153-158.

88. Feinberg EJ, Steinberg WM, Banks BL, et al. How long to abstain from eating red meat before fecal occult blood tests. Annals of Internal Medicine 1990;113: 403-404

89. Fludger S, Turner AM, Harvey RF, et al. Controlled prospective study of faecal occult blood screening for colorectal cancer in Bury, black pudding capital of the world. *British Medical Journal* 2002;325: 1444-1445.

90. Sinatra MA, St John DJB and Young GP. Interference of plant peroxidases with guaiac- based fecal occult blood tests is avoidable. *Clinical Chemistry* 1999;45: 123-126.

91. Robinson MH, Pye G, Thomas WM, et al. Haemoccult screening for colorectal cancer: the effect of dietary restriction on compliance. *European Journal of Surgical Oncology* 1994;20: 545-548.

92. Pignone M, Campbell MK, Carr C, et al. Meta-analysis of dietary restriction during faecal occult blood testing. *Effective Clinical Practice* 2001;4: 150-156.

93. Konrad G. Dietary interventions for fecal occult blood test screening: systematic review of the literature. *Canadian Family Physician* 2010;56: 229-238.

94. Carroll MRR, Seaman HE and Halloran SP. Tests and investigations for colorectal cancer. *Clinical Biochemistry* 2014;47: 921–939.

95. Clarke P, Jack F, Carey F, et al. Medications with anticoagulant properties increase the likelihood of a negative colonoscopy in faecal occult blood test population screening. *Colorectal Disease* 2005;8: 389-392.

96. Konrad G and Katz, A. Are medication restrictions before FOBT necessary? Practical advice based on a systematic review of the literature. *Canadian Family Physician* 2012;58: 939-948.

97. Kahi CJ, Anderson JC and Rex DK. Screening and surveillance for colorectal cancer: state of the art. *Gastrointestinal Endoscopy* 2013;77: 335-350

98. Logan RFA, Little J, Hawtin P G, et al. Effect of aspirin and non-steroidal anti inflammatory drugs on colorectal adenomas: case-control study of subjects participating in the Nottingham faecal occult blood screening programme. *British Medical Journal* 1993;307: 285-289.

99. Rothwell PM, Wilson M, Elwin C, et al. Long-term effect of aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised trials. *The Lancet* 2010;376: 1741-1750.

100. Fraser CG. Faecal occult blood tests: Life savers or outdated colorectal screening tools? *Clinical Laboratory News* 2011;37: 8-10.

101. U.S. Preventive Services Task Force. Screening for Colorectal Cancer: Recommendation and Rationale. *Annals of Internal Medicine.* 2002;137: 129-131.

102. Hundt S, Huag U and Brenner H. Comparative evaluation of immunochemical fecal occult blood tests for colorectal adenoma detection. *Annals of Internal Medicine* 2009;150: 162-169.

103. Brenner H, Haug U and Hundt S. Inter-test agreement and quantitative cross-validation of immunochromatographical fecal occult blood test. *International Journal of Cancer* 2010;127: 1643-649.

104. Allison JE, Fraser CG, Halloran SP, et al. Comparing faecal immunochemical tests: improved standardization is needed *Gastroenterology* 2012;142: 422–431

105. Fraser, CG. Allison, JE. Halloran, SP, et al.on behalf of the Expert Working Group on fecal immunochemical tests for hemoglobin, Colorectal Cancer Screening Committee, World Endoscopy Organization. A proposal to standardize reporting units for fecal immunochemical tests for hemoglobin. *Journal of the National Cancer Institute* 2012;104: 810-814.

106. Tinmouth J, Lansdorp-Vogelaar I and Allison JE. Fecal immunochemical tests versus guaiac faecal occult blood tests: what clinicians and colorectal cancer screening programme organisers need to know. *Gut* 2015;64: 1327-1337.

107. Moss SM, Campbell C, Mellia J, et al. Performance measures in three rounds of the English bowel cancer screening pilot. *Gut* 2011;61: 101-107.

108. National Services Scotland. Infomation Services Division, Bowel Screening Programme annual update. http://www.isdscotland.org/Publications/index.asp [Accessed: 12 October 2015]

109. Hoffman RM, Steel S, Yee EFT, et al. Colorectal cancer screening adherence is higher with fecal immunochemical tests than guaiac-based fecal occult blood tests: a randomized controlled trial. *Preventative Medicine* 2010;50: 297-299.

110. Oort FA, van Turenhout ST, Coupe VMH, et al. Double sampling of a faecal immunochemical test is not superior to single sampling for detection of colorectal neoplasia: a colonoscopy controlled prospective cohort study. *BMS Cancer* 2011;11: 434.

111. Rozen P, Comaneshter D, Levi Z, et al. Cumulative evaluation of a quantitative immunochemical fecal occult blood test to determine its optimal clinical use. *Cancer* 2010;116: 2115-2125.

112. Terhaar sive Droste SJ, Oort FA, van der Hulst RWM, et al. Higher fecal immunochemical test cutoff levels: lower positivity rates but still acceptable detection rates for early-stage colorectal cancers. *Cancer Epidemiology, Biomarkers and Prevention* 2011;20: 272-280.

113. Brown LF and Fraser CG. Effect of delay in sampling on haemoglobin determined by faecal immunochemical tests. *Annals of Clinical Biochemistry* 2008;45: 604-605.

114. van Roon AHC, Hol L, van Vuuren AJ, et al. Are fecal immunochemical test characteristics influenced by sample return time? A population-based colorectal cancer screening trial. *The American Journal of Gastroenterology* 2011;107: 99–107

115. van Rossum LGM, van Rijn AF, van Oijen GH, et al. False negative fecal occult blood tests due to delayed sample return in colorectal cancer screening. *International Journal of Cancer* 2009;125: 746-50.

116. Grazzini G, Ventura L, Zappa M, et al. Influence of seasonal variations in ambient temperatures on performance of immunochemical faecal occult blood test for colorectal cancer screening: observational study from the Florence district. *Gut* 2010;59: 1511-1515.

117. Lee CS, O'Gorman P, Walsh P, et al. Immunochemical faecal occult blood tests have superior stability and analytical performance characteristics over guaiac-based tests in a controlled in vitro study. *Journal of Clinical Pathology* 2011;64: 524-528.

118. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. *Gastroenterology* 2008;135: 82-90.

119. Guittet L, Bouvier V, Mariotte N, et al. Camparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. *Gut* 2007;56: 210-214.

120. Rabeneck L, Rumble RB, Thompson F, et al. Fecal immunochemical tests compared with guaiac fecal occult blood tests for population-based colorectal cancer screening. *Canadian Journal of Gastroenterology* 2012;26: 131-147.

121. Vilkin A, Rozen P, Levi Z, et al. Performance characteristics and evaluation of an automated-developed and quantitative, immunochemical, fecal occult blood screening test. *American Journal of Gastroenterology* 2005;100: 2519–2525.

122. Grazzini G, Visioli CB, Zorzi M, et al. Immunochemical faecal occult blood test: number of samples and positivity cutoff. What is the best strategy for colorectal cancer screening? *British Journal of Cancer* 2009;100: 259-265.

123. Kapidzic A, Grobbee EJ, Hol L, et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. *American Journal of Gastroenterology* 2014;109: 1257-1264.

124. van Roon AHC, Goede SL, van Ballegooijen M, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening *Gut* 2012;62: 409-415.

125. Wilschut JA, Habbema JDF, van Leerdam ME, et al. Fecal occult blood testing when colonoscopy capacity is limited. *Journal of the National Cancer Institute* 2011;103: 1741-1751.

126. Brenner H, Tao S and Huag U. Low-dose aspirin use and performance of immunochemical fecal occult blood tests. *Journal of the American Medical Association* 2010;304: 2513-2520.

127. Levi Z, Rozen P, Hazazi R, et al. Sensitivity, but not specificity, of a quantitative immunochemical fecal occult blood test for neoplasia is slightly increased by the use of low-dose aspirin, NSAIDs, and anticoagulants. *American Journal of Gastroenterology* 2009;104: 933-938

128. Kahi CJ and Imperiale TF. Do aspirin and nonsteroidal anti-inflammatory drugs cause false-positive fecal occult blood test results? A prospective study in a cohort of veterans. *The American Journal of Medicine* 2004;117: 837-841.

129. Soares-Weiser K, Burch J, Duffy S, et al. Diagnostic accuracy and cost-effectiveness of faecal occult blood tests (FOBT) used in screening for colorectal cancer: a systematic Review. *Centre for Reviews and Dissemination, York.* 2007.

130. Berchi C, Guittett L, Bouvier V, et al. Cost-effectiveness analysis of the optimal threshold of an automated immunochemical test for colorectal cancer screening. *International Journal of Technology Assessment in Health Care* 2010;26: 48-53.

131. Guildford Medical Device Evaluation Centre (GMEC). Research in Bowel Screening. NHS Bowel Cancer Screening Programme. http://www.worldendo.org/fit-ewg-publications.html [Accessed: 30 March 2014]

132. CPA. Clinical Pathology Accreditation (UK) Ltd. http://www.cpa-uk.co.uk/ [Accessed: 04 June 2014]

133. Programme Resources: Bowel Screening Standards March 2015. Healthcare Improvement Scotland. http://www.healthcareimprovementscotland.org/our_work/cancer_care_improvement/programme_resources/bowel_screening_standards.asp [Accessed: 20 April 2015]

134. WHO International Standard Haemiglobincyanide NIBSC code: 98/708 . National Institute for Biological Standards and Control http://www.nibsc.org/products/brm_product_catalogue/detail_page.aspx?catid=98/708 [Accessed: 20 April 2015]

135. Petersen PH and Fraser CG. Strategies to set global analytical quality specifications in laboratory medicine: 10 years on from the Stockholm consensus conference *Accreditation and Quality Assurance* 2010;15: 323-330.

136. The retention and storage of pathological records and specimens (5th edition) . Royal College of Pathologists. http://www.rcpath.org/publications-media/publications/ [Accessed: 20 April 2015]

137. van Rossum LG, van Rijn AF, Laheij RJ, et al. Cutoff value determines the performance of a semi-quantitative immunochemical faecal occult blood test in a colorectal cancer screening programme. *Bristish Journal of Cancer* 2009;101: 1274-1281.

138. Bossuyt PM, Reitsma JB, Bruns DE, et al Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clinical Chemistry* 2003;49: 1-6.

139 Hanley J and McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982:143; 29-36

140. Fraser CG, Mathew CM, McKay K, et al. Automated immunochemical quantitation of haemoglobin in faeces collected on cards for screening for colorectal cancer. *Gut* 2008;57: 1256-1260.

141. Symonds EL, Osborne JM, Cole SR, et al. Factors affecting faecal immunochemical test positive rates: demographic, pathological, behavioural and environmental variables. *Journal of Medical Screening* 2015 doi:10.1177/0969141315584783 [accessed 17 May 2015]

142. Gnatta E, Zaninotto M, Epifani MG, et al. A new sampling device for faecal immunochemical testing: haemoglobin stability is still an open issue*. Clinical Chemistry and Laboratory Medicine* 2014;52: 1203-1209.

143. Dancourt V, Lejeune C, Lepage C, et al. Immunological faecal occult blood tests are superior to guaiac-based tests for the detection of colorectal neoplasms. *European Journal of Cancer* 2008;4: 2254-2258.

144. National Services Scotland. Health and Wellbeing Profiles 2010 Scotland overview report. Scottish Public Health Observatory.

http://www.scotpho.org.uk/web/FILES/Profiles/2010/ScotPHO_overview_2010-tiffs-print-and-web.pdf [Accessed: 12 June 2013].

145. Steele RJC, Kostouru I, McClements P, et al. Effects of age, gender and deprivation on key performance indicators in a FOBT based colorectal cancer screening programme. *Journal of Medical Screening* 2010;17: 68-74

146. Cole SR, Young GP, Esterman A, et al. A randomized trial of the impact of new fecal hemoglobin test technologies on population participation in screening for colorectal cancer. *Journal of Medical Screening* 2003;10: 117-122.

147 Hol L, Wilschut JA, van Ballegooijen M, et al. Screening for colorectal cancer: random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. *British Journal of Cancer* 2009;100: 1103-1110.

148. Hol L, Leerdam ME, Ballegooijen M, et al.Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. *Gut* 2010;59: 62-68.

149. Allison JE, Sakoda LC, Levin TR, et al. Screening for colorectal neoplasms with new fecal occult blood tests: update on performance characteristics. *Journal of the National Cancer Istitute* 2007;99: 1462-1470.

150. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. *American Journal of Gastroenterology* 2010;105: 2017-2025.

151. Fraser CG, Rubeca T, Rapi S, et al. Feacal haemoglobin concentrations vary with sex and age, but are not transferable across geography for colorectal cancer screening. *Clinical Chemistry and Laboratory* Medicine 2014;52: 1211-1216.

152. Chaing TH, Chuang SL, Chen SL, et al. Difference in performance of fecal immunochemical tests with the same haemoglobin cut-off concentration in a nationwide colorectal cancer screening program. *Gastroenterology* 2014;147: 1317-1326.

153. Lee TJW, Clifford GM, Rajasekhar P, et al. High yield of colorectal neoplasia detected by colonosopy following a positive faecal occult blood test in the NHS Bowel cancer Screening Programme. *Journal of Medical Screening* 2011;15: 82-86.

154. Goulard H, Boussac-Zarebska M, Ancelle-Park R, et al. French colorectal cancer screening pilot programme results of the first round. *Journal of Medical Screening* 2008;15: 143-8.

155. Fraser CG. Screening for colorectal neoplasia with faecal tests. *Lancet Oncology* 2011;12: 516-517.

156. About CLSI. Clinical and Laboratory Standards Institute. http://clsi.org/about-clsi/ [Accessed: 20 April 2015].

157. Horowtiz GL, et al. Clinical Laboratory Standards Institute. Defining, establishing, and verifying reference intervals in the clinical laboratory - approved guideline third edition. (2008).Wayne, PA.

158. Chen JS, Yen AMF, Chiu SHY, et al. Baseline immunochemical faecal occult blood test (iFOBT) concentration as a predictor of incident colorectal neoplasia: longitudinal follow-up of a Taiwanese population-based colorectal cancer screening cohort. *Lancet Oncology* 2011;12: 551-558.

159. Brenner H, Altenhofen L and Hoffmeister M. Sex, age and birth cohort effects in colorectal neoplasms. *Annals of Internal Medicine* 2010;152: 697-703.

160. Rozen P, Levi Z, Hazazi R, et al. Identification of colorectal adenomas by a quantitative immunochemical fecal occult blood screening test depends on adenoma characteristics, development threshold used and number of tests performed. *Alimentary Pharmacology & Therapeutics* 2009;29: 906-917.

161. Levi Z, Rozen P, Hazazi R, et al. A quantitative immunochemical fecal occult blood test for colorectal neoplasia. *Annals of Internal Medicine* 2007;146: 244-255.

162. Rocky D. Occult and obscure gastrointestinal bleeding: causes and clinical management. *Nature Reviews, Gastroenterology and Hepatology* 2010;7: 265-279.

163. Khalid-de-Bakker CA, Jonkers DM, Sanduleanu S, et al. Test performance of fecal occult blood testing and sigmoidoscopy compared with colonoscopy screening for colorectal advanced adenoma. *Cancer Prevention Research* 2011;4: 1563-1571.

164. Yeoh KG, Ho KY and Chiu HM. The Asia-Pacific colorectal screening score: a validated tool that stratifies risk for colorectal advanced neoplasia in asymptomatic Asian subjects *Gut* 2011;60: 1236-1241.

165. Marshall T, Lancashire R, Sharp D, et al. The diagnostic performance of scoring systems to identify symptomatic colorectal cancer compared to current referral guidance. *Gut* 2011;60: 1242-1248.

166. Manz M, Burri E, Rothen C, et al. Value of fecal calprotectin in the evaluation of patients with abdominal discomfort: an observational study. *BMC Gastroenterology* 2012;12: 5.

167. Fraser CG. Faecal occult blood tests - eliminate, enhance or update? *Annals of Clincial Biochemistry* 2008;45: 117-121.

168. Young GP, Fraser CG, Halloran SP, et al. Guaiac-based faecal occult blood testing for colorectal cancer screening - an obsolete strategy? *Gut* 2012;61: 959-960.

169. Auge JM, Fraser CG, Cristina R, et al. Clinical utility of one v two FIT samples in the detection of advanced colorectal neoplasia in symptomatic patients. *Clinical Chemistry and Laboratory Medicine* 2015 doi: 10.1515/cclm-2015-0388

170. Rodriguez-Alonso L, Rodriguez-Moranta F, Ruiz-Cerulla A, et al. An urgent referral strategy for symptomatic patients with suspected colorectal cancer based on a quantitative immunochemical faecal occult blood test. *Digestive and Liver Disease* 2015 doi 10.1016/j.dld.2015.05.004. [Accessed 17 June 2015]

171. van Dam L, Kuipers EJ and van Leerdam ME. Performance improvements of stool based screening tests. *Best Practice & Research Clinical Gastroenterology* 2010;24: 479-492.

172. Allison JE, Fraser CG, Halloran SP, et al. Comparing fecal immunochemical tests: improved standardization is needed. *Gastroenterology* 2012;142: 422-424.

173. Omata F, Shintani A, Isozaki M, et al.Diagnostic performance of quantitative fecal immunochemical test with multivariate prediction model for colorectal neoplasms in asymptomatic individuals. *European Journal of Gastroenterology and Hepatology* 2011;23: 1036-1041.

174. Hippisely-Cox J and Coupland C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. *British Journal of General Practice* 2012;62: e29-e37.

175. Collins GS and Altman DG. Identifying patients with undetected colorectal cancer, an independent validation of QCancer. (Colorectal). *British Journal of Cancer* 2012;107: 260-265.

176. Elsafi SH, Alqahtani NI, Zakary NY, et al. The sensitivity, specificity, predictive values, and likelihood ratios of fecal occult blood test for the detection of colorectal cancer in hospital settings. *Clinical and Experimental Gastroenterology* 2015;8: 279–228.

177. Wong MCS, Lam TYT, Tsoi KKF, et al. A validated tool to predict colorectal neoplasia and inform screening choice for asymptomatic subjects. *Gut* 2014;64: 776-786.

178. Digby J, Fraser CG, Carey FA, et al. Faecal haemoglobin concentration is related to severity of colorectal neoplasia. *Journal of Clinical Pathology* 2013;66: 415-419

179. Cubiella J, Salve M, Diaz-Ondina M, et al. Diagnostic accuracy of the faecal immunochemical test for colorectal cancer in symptomatic patients: comparison with NICE and SIGN referral criteria. *Colorectal Disease* 2014;16: 273-282.

180. Wells BJ, Kattan MW, Cooper G, et al. ColoRectal cancer predicted risk online (CRC-PRO) calculator using data from the Multi Ethnic cohort study. *Journal of the American Board of Family Medicine* 2014;27: 42-56.

181. Garcia M, Binefa G, Benito L, et al. Fecal haemoglobin concentration as a measurement of risk to tailor colorectal cancer screening. Are we there yet? *European Journal of Cancer Prevention* 2015; 24:321–327

182. Chen G, Mao B, Pan Q, et al. Prediction rule for estimating advanced colorectal neoplasm risk in average-risk populations in southern Jaingsu Province. *Chinese Journal of Cancer Research* 2014; 26: 4-11.

183. Wang JY, Li ZT, Zhu YM, et al. Utility of the Asia-Pacific colorectal screening scoring system and the presence of metabolic syndrome components in screening for sporadic colorectal cancer. *World Journal of Gastroenterology* 2014; 20: 11394-11399.

APPENDICES

APPENDIX 1 STANDARD OPERATING PROCEDURE 31 HOW TO OPERATE OC-SENSOR

Department of Biochemical Medicine
NHS Tayside

Title	Quantitative FIT using OC Sensor

Document Reference.Version	TSCRPU 34.1
CPA Standard	F2
Author	P McDonald
Named Reviewer	P McDonald
Approved by Signature Date of Issue	J Strachan 28 July 2010
Review Frequency	Annual
Copy	1 of 2
Location of Copies	1. Master 2. Q Pulse 3. SBSL

References of Documents Cited Herein*	TSCRPU 3 TSCRPU 4 TSCRPU 6 TSCRPU 8 TSCRPU 10 TSCRPU 20 TSAFE 10 NSAFE 52 TSCRPU 25 TSCRPU 34 F1.1 Kit insert Simple Operation Manual Manufacturers Performance Data

*: These documents must be read in conjunction with this document.

Clinical Relevance & Principle of Examination

The aim of the Scottish Bowel Screening Programme is to screen persons between the ages 50-74, throughout Scotland, for faecal occult blood using a two-year recall. Quantitative FIT is being used as a trial in NHS Tayside and Ayrshire and Arran from July 2010. All faecal occult blood test positive participants are referred to colonoscopy for further investigation.

Measurement using the OC Sensor DIANA is based on latex agglutination. An antigen-antibody reaction is a specific reaction that occurs between an antigenic determinant and the active group of an antibody. The amount of bonding depends on the concentration of antigen or antibody.

A latex agglutination reaction is the clumping of antigen or antibody sensitized polystyrene latex particles through an antigen-antibody reaction. A light beam is passed through the reaction liquid to measure the changes in the intensity of the transmitted light beam (latex turbidimetry), and changes in the intensity of the scattered light beam (latex nephelometry). The OC Sensor DIANA uses latex turbidimetry to measure the amount of haemoglobin present in stool samples via the sampling device.

References

Refer to Instructions for Use

Safety Data

Biochemical Medicine Health and Safety Policy Health and Safety Codes of Practice ***TSAFE 10***, ***NSAFE 52***, and ***TSCRPU 25***
All staff must wear gloves and laboratory coats whilst testing due to potential biological and chemical hazards.

Equipment and Specialist Supplies

2 x OC Sensor DIANA, Eiken Chemical Co., Ltd

Domestic bleach
Purified water
100-1000 µl pipette and tips
Sample cups

Zebra printer, labels

Reagents, Standards or Calibrants, and Internal Quality Control Materials

OC Diana Latex REF V-PZ01, Eiken Chemical Co., Ltd. Supplied ready to use, stored at 2-10^{o}C. These are kept in room CI-007, reagent store, in fridge 2.

OC Diana Buffer REF V-PZ03, Eiken Chemical Co., Ltd. Supplied ready to use, stored at 2-10^{o}C. These are kept in room CI-007, reagent store, in fridge 2.

Standards and Calibrants

OC Standard REF V-IX50, Eiken Chemical Co., Ltd. This is stored at 2-10^{o}C unopened for up to 1 year. These are kept in room CI-007 reagent store in fridge 2.

The standard must be reconstituted by adding 1.0ml purified water and mixing gently by inversion several times and leave for a minimum of 15 minutes. The diluent is ready made. Using the pre programmed calibration rack, 2ml of diluent is placed in position 1 and 250 µl of low control and 250 µl of high control in positions 9 and 10 respectively.

Internal Quality Control Materials

There are two IQC OC-Control Low REF V-PH57 and OC-Control High REF V-PH58, Eiken Chemical Co., Ltd. These are supplied freeze-dried and are stable at 2-10^{o}C for 1 year. These are kept in room CI-007, reagent store, in fridge 2.

Reconstitute by adding 1.0ml purified water and mix gently by inversion several times and leave for a minimum of 15 minutes. When using a new vial aliquot 200 µl portions into eppindorf tubes and store at –20^{0}C immediately. Using the pre programmed calibration rack 250 µl of low control and 250 µl of high control are put in positions 9 and 10 respectively.

Specimen Requirements and Means of Identification

Participants who receive a FIT kit receive an invitation to test. In with the letter is a label containing the participants name, kit number CHI number and a space to write the sample date. This must be taken from the letter by the participant and attached to the leak-proof bag that holds the tube supplied in the invitation pack.

The participant receives an instruction leaflet detailing the sample collection method.

The participant must write the sample date on the label. The samples are stable for 10 days after application of faeces.

Samples are tested the day they arrive in the laboratory. If this does not happen, the samples can be stored at 4^{o}C overnight in Fridge 2. Once tested the samples are stored at –20^{o}C for one week to allow results letters to be sent to participants and any queries to be dealt with. Samples are identified using a 10-digit numeric barcode label

Method

Calibration

Calibration is performed once a month or whenever the Lot of latex changes. The calibration method is a six point nonlinear calibration using OC Standard REF V-IX50.

Refer to Simple Operation Manual

Performance of Examination

Kits that arrive in the laboratory are Booked-in and Sorted, see ***TSCRPU 3***.

Proceed with testing of non-problematic kits as follows, for all others see ***TSCRPU 6.***

<u>Start-up Analyser</u>

a) Using the maintenance schedule TSCRPU 34 F1.1, ensure that the drain tank is empty, the purified water, wash solution and buffer are full. If these are topped up prime them, see section e
b) The IQC and latex are removed from the fridge in preparation for putting on the machine. Prepare the IQC, during the fifteen minutes they reconstitute turn the OC sensor on using the button at the front right hand side of the machine.
c) The front cover of the analyser can be lifted when the MENU screen is displayed
d) Place the latex on the machine. The lids from the latex should be removed and placed alongside the lid for the buffer on the front central section of the machine

The cover on the latex positions can then be lifted and two bottles of latex inserted. The latex cover must then be replaced followed by the large front cover

e) Prime the machine if necessary – that is if purified water, wash or buffer have been topped up or replaced
   a. To prime the machine on the touch screen press SUPPORT, then PREP FUNCTIONS/SUPPORT
   b. Then select PRIME from tabs
   c. Select NORMAL PRIME and select YES for any of the items that have been changed
   d. Press START, when the prime is completed return to the main MENU
   e. If buffer has been replaced go to LATEX and RESET buffer.

f) The IQC can now be dispensed into cups for testing. Using a IQC rack (dark blue in colour) place two clear cups in the IQC positions at the far right of the rack. Using a pipette measure out 200µl of the low IQC and dispense into the cup on the left, repeat with the high IQC dispensing into the cup on the right.

g) Aliquots of IQC can be stored in the freezer for up to five days. They must have the type of control and date they were frozen clearly written on them.

h) Place the rack on the tray on the right hand side of the machine

i) Check that all the fluids are topped up, the lids are off the latex and that nothing is obscuring the path of the rack. On the screen press TEST then START, the machine performs a self-check before starting. If the self check fails get a senior to look at the problem

j) When the IQC results are printed out they must be checked against the local ranges for the current batch. If they are within the range they are accepted. If they are outside the range troubleshooting is performed in the following order - check the fluids, rerun the controls, make up new controls or recalibrate the machine.

<u>Sample Runs</u>

a) Log onto PC
b) Ensure that the Zebra printer is selected as the default printer
c) Click on the barcode icon
d) Pick up a box of kits that have been booked and sorted and take it through to the PC. Starting at the front of the box where the earliest kits are remove the first kit
e) Scan the kit label, the barcode should be read and a label printed automatically.
f) If the label does not print automatically it should be batched and kept to the end of a run, the number can be typed in manually, the label will print automatically when the 10th digit is entered.
g) The label should be checked against the number on the bag. NB. If the labels do not match then reprint the label. When the labels match, the sample tube should be removed from the bag and the new label attached to the tube
h) The sample bags must be kept on order in batches with the rack number written on the batch.
i) The tube can now be loaded into a light blue sample rack. Repeat process until 10 tubes are loaded into the rack, or all the labels have been completed
j) Place the rack on the tray at the right hand side of the machine
k) On the screen press TEST then START
l) If more samples are to be set running on the machine press SET SAMPLE and then START
m) Once testing of the first 100 samples has been completed another set of IQC should be run and accepted. This process continues until that days samples are complete.
n) Samples are stored and disposed as per ***TSCRPU 10***

<u>Recording and Calculation of Results</u>

IQC results and patient results are sent in separate reports.

To send IQC reports.

a) The analyser must be finished before exporting data.
b) Ensure that the PC is switched on, the printer is set to HP DeskJet 1600 and the MAST Data Integration Software (DIS) is open.
c) From the analyser menu select SUPPORT
d) Select PROCESS IQC
e) Highlight the type of IQC - QC 1 for low control results and QC2 for high control results, and then select the current LOT.
f) Using the tab select INTRA-DAY/INTER-DAY
g) Highlight the data you want to send, press SELECT S (to start) and SELECT to finish
h) When you are ready to export the data press OUTPUT then EXT MEDIA, then START.
i) When the second set of results are set there is an option to ADD or DELETE, press ADD
j) The data has now been sent to DIS, to close QC menu press CONTINUE, REGISTER.

At the PC check the QC results for consistency, verify the results and print them.
IQC results must be verified before a run of samples are tested.

To send participant results.

k) The analyser must be finished before exporting data.
l) Ensure that the PC is switched on, the printer is set to HP Deskjet 1600 and the MAST Data Integration Software (DIS) is open.
m) From the analyser menu select SUPPORT
n) Select PROCESS DATA
o) The most recent sample run is located at the top left hand corner of the screen. It is identified by the date and time it happened.
p) Highlight the data you want to send, press SELECT S (to start) and SELECT to finish. This selects a complete run.
q) Once isolated, highlight the individual test data you wish to export and select TEST DATA.
r) All the results from the run now appear press SELECT S (to start) press the bottom result visible (all the results visible will turn blue) scroll down once, press the bottom result so that all results turn blue repeat this until the whole run is selected
s) When you are ready to export the data press OUTPUT then EXT MEDIA, then START.
t) The data has now been sent to DIS, to close SUPPORT menu press CONTINUE, REGISTER.

Log into the MAST DIS. At the PC check the participant results for consistency –ensuring there is a variety of results and no trends that would indicate a problem within the run, then press verification complete for that set of results and print them.

These printed results must be checked against the analyser printout to ensure that there has not been any error during transmission of results. The analyser printout should be ticked for each result that is correctly matched. If any results do not match it must be brought to the attention of a senior immediately

Two Screeners enter results into BoSS from the printed work list. One person reads from the printed results sheet, the second person types the kit number into BoSS before entering the result. Both people ensure the correct result is entered then proceed to the next participant result. When entering the result, 0 is entered for a negative result, 3 is entered for a positive result. All results printouts are kept with the original analyser data stapled to them.

Maintenance is carried out as per the schedule attached (TSCRPU 34 F1.1)

All maintenance procedures are found in the CLOSE MODE, MAINTENANCE, PROCEDURE screens on the analyser.

Limitations of the Procedure, Most Frequent Problems and Troubleshooting

Almost no reaction occurs with non-human haemoglobin.

Almost no effect on the measurement value was found from bilirubin (25mg/dl), lipids (0.6% intra lipid), ascorbic acid (40 mg/ml), protein (2.5 g/dl bovine serum albumin), glucose (4.0 g/dl) or barium sulphate (25 mg/dl).

A senior member of staff must perform troubleshooting. If there is a communication error, shut down the machine and start it up again. If the controls are out of the limits, make a new batch and repeat them. If there is a mechanical fault (rack not ejecting or stuck rack) stop the run and remove the obstruction.

Advice on trouble shooting problems can be sought from Iain McElarney on 07855 416564, office 0151 933 7277 or imcelarney@mastgrp.com.

Range of Expected Results

The OC Sensor DIANA, Eiken Chemical Co., Ltd is calibrated to measure 0 – 1000 ng Hb/ml of buffer, with a prozone check function above 1000 ng Hb/ml buffer

Validation and Reporting

The control values are checked against the locally set range obtained after analysing QC1low and QC high ten times. Data attached.
Screeners enter results into the BoSS system. A senior Screener or Biomedical Scientist can change the result that same day. At 9 pm the results are moved to mailer files for printing result letters to participants.

Reference Intervals

A result is reported as positive when the amount of haemoglobin measured is over 400 ng/ml buffer. This will be a training data set and will be changed to give a 2% positivity rate.

Performance Criteria

BETWEEN RUN PRECISION

OC 2

MEAN	148.7	623.9
SD	7.7	28.6
CV(%)	5.2	4.6

OC 1

MEAN	153.9	616.9
SD	6.1	23.1
CV(%)	3.9	3.7

WITHIN-RUN PRECISION

	SAMPLE 1	SAMPLE 2	SAMPLE 3	SAMPLE 4	LOW	HIGH
MEAN	25.8	47.1	72.4	83.5	172.1	708.4
SD	6.5	8.1	3.9	6.1	5.8	21.1
CV(%)	25.2	17.2	5.4	7.3	3.4	3.0

Two-tailed distribution, two sample equal variance

T test	Low 0.1100	High 0.5548

See also Manufacturers Performance Data

Disposal of Materials

See ***TSCRPU 10***

ANALYSER	ANALYSIS DATE	ANALYSIS T	RACK NO#	POSITION	SAMPLE ID	CC NO#	Hb ng/ml	JUDGEMENT	ERROR CODE	
DIANA01	30/12/2010	11:03	4	1	5115369341	6	0	-		
DIANA01	30/12/2010	11:03	4	2	8715340598	6	3	-		
DIANA01	30/12/2010	11:03	4	3	6515369387	6	0	-		
DIANA01	30/12/2010	11:03	4	4	4115364714	6	0	-		
DIANA01	30/12/2010	11:04	4	5	5215367109	6	28	-		
DIANA01	30/12/2010	11:04	4	6	2315369392	6	0	-		
DIANA01	30/12/2010	11:04	4	7	4315366865	6	0	-		
DIANA01	30/12/2010	11:04	4	8	1115362299	6	0	-		
DIANA01	30/12/2010	11:04	4	9	1315259807	6	15	-		
DIANA01	30/12/2010	11:05	4	10	5615362087	6	496	+		
DIANA01	30/12/2010	11:05	6	1	8415274403	6	0	-		
DIANA01	30/12/2010	11:05	6	2	4615327598	6	0	-		
DIANA01	30/12/2010	11:05	6	3	5715349019	6	0	-		
DIANA01	30/12/2010	11:05	6	4	3315367089	6	0	-		
DIANA01	30/12/2010	11:06	6	5	1315247963	6	0	-		
DIANA01	30/12/2010	11:06	6	6	6715333675	6	217	-		
DIANA01	30/12/2010	11:06	6	7	8715338330	6	0	-		
DIANA01	30/12/2010	11:06	6	8	5515348884	6	0	-		
DIANA01	30/12/2010	11:07	6	9	2315366998	6	14	-		
DIANA01	30/12/2010	11:07	6	10	5615274328	6	0	-		
DIANA01	30/12/2010	11:07	9	1	3115335895	6	0	-		
DIANA01	30/12/2010	11:07	9	2	7115367003	6	0	-		
DIANA01	30/12/2010	11:07	9	3	9515333372	6	91	-		
DIANA01	30/12/2010	11:08	9	4	3115375396	6	0	-		
DIANA01	30/12/2010	11:08	9	5	8415358319	6	0	-		
DIANA01	30/12/2010	11:08	9	6	3115349062	6	0	-		
DIANA01	30/12/2010	11:08	9	7	7115369397	6	0	-		
DIANA01	30/12/2010	11:08	9	8	9315342813	6	0	-		
DIANA01	30/12/2010	12:31	15	1	3715397662	6	6	-		

ANALYSER	DATA TYPE	ANALYSIS DATE	ANALYSIS T	RACK NO#	POSITION I	QC LOT	CC NO#	Hb ng/ml	JUDGEMENT	ERROR CODE	
DIANA01	C1	30/12/2010	08:47	92	9	9013	6	158	✓		
DIANA01	C1	30/12/2010	10:06	92	9	9013	6	158	✓		
DIANA01	C2	30/12/2010	08:47	92	10	9013	6	685	✓		
DIANA01	C2	30/12/2010	10:06	92	10	9013	6	692	✓		

acceptable RTM
30/12/10

APPENDIX 3 FFLT LABORATORY REPORT

FIT as a First line Test Evaluation – Overview of Laboratory Phase

The "FIT as a First line Test" evaluation (FFLT) within the Scottish Bowel Screening Programme got underway on 01 July 2010. All those eligible participants within NHS Tayside and NHS Ayrshire & Arran were sent one Eiken Chemical Co FIT specimen collection tube (SFIT) rather than the usual *hema-screen* gFOBT. The Laboratory phase has ended and data from this are being collected. The data in this report represent final outcomes.

1. Number of invitations sent out and tubes returned.

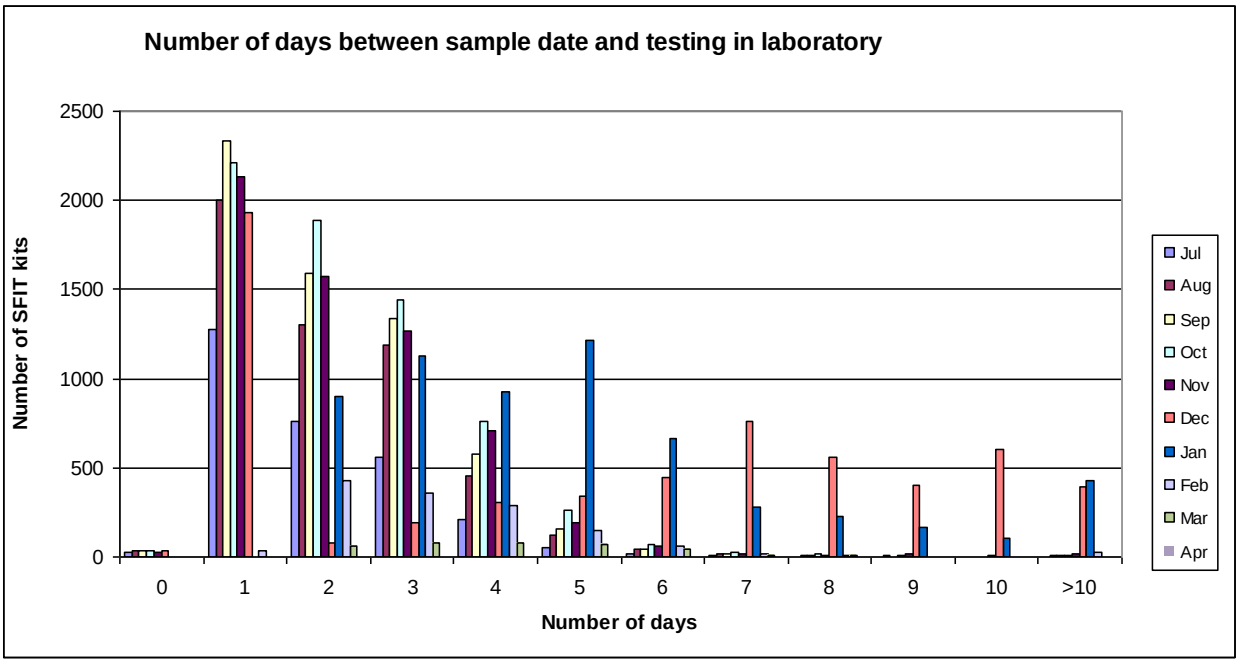
SFIT kits ceased being sent to participants on 12 January 2011. As of 12 January 2011, the number of SFIT kits [flow-pack with tube, return envelope and zip-lock bag, information on how to do the test and the Know the Facts leaflet] sent out from the Centre as initial invitations was 66225.. The number of tubes that have been received in the Laboratory as of 12 April 2011 was 40125. Thus, initial calculations indicate overall response to the invitation to participate in the study was 60.6%. Of the total number of samples received in the laboratory 1405 were initially Untestable and the participant was sent a second sampling device in order to complete their testing cycle. The number of participants who completed their cycle with a positive or negative result was 38720; this is 58.5% of the invited population.

2. Number of days between sample being dated by participant and result entry in Laboratory.

The graph below shows that 98% of SFIT tubes were returned to the Laboratory within 10 days or less from the date written on the label by the participant (this includes time spent in the post).

The graph also shows that in December and January, the adverse weather and its consequences for the post caused the number of expired SFIT to rise dramatically. This pattern was not replicated in other months

Graph 1 Data for whole study return of SFIT kits



3. Turnaround of SFIT analyses in Centre Laboratory

Table 1. Data for complete study

Calendar Days to Report Result		SFIT	Cumulative %
1		37438	93.3
2		1858	95.4
3		308	99.2
4		215	99.5
5		197	99.5
6		78	99.8
7		29	99.9
8		2	100
Total		40125	

These data show that more than 95% of SFIT kits can be assayed in the Laboratory within 2 days in normal circumstances. The current QIS standard demands that 95% of kits are tested in the laboratory within 5 working days (7 calendar days) of receipt.

4. FFLT samples returned by Month

Table 2. Data for complete study

	Males		**Females**	
	Samples	Cumulative	Samples	Cumulative
July 2010	1285	1285	1643	1643
August 2010	2513	3798	2764	4407
September 2010	2825	6623	3316	7723
October 2010	3146	9769	3598	11321
November 2010	2842	12611	3216	14537
December 2010	1851	14462	1893	16430
January 2011	2797	17259	3340	19770
February 2011	621	17880	714	20484
March /April 2011	178	18058	178	20662

5. Laboratory outcome data

The numbers of unlabelled, undated, and spoiled SFIT kits are proportionately lower than for gFOBT. The number of SFIT kits with mechanical damage (reported as non technical fail with the participants being sent another SFIT kit) is noticeably higher than for gFOBT. There have been no manufacturer related technical fails of SFIT. Note that there were a significantly higher number of “expired” kits in December, simply due to the non-delivery of mail within the usual time frame caused by adverse weather and consequent transport problems. All participants who have encountered a pre analytical problem are sent a repeat sampling device.

Table 3. Laboratory outcome data for the complete study

ITEM	SFIT		
			%
1	**Invitations**	66225	
	Testable SFIT		
2	Positive	943	2.4
3	Negative	37777	97.6
4	Total tested	38720	
5	Total tested/Invitations		58.5
6	Total tested/Tested and Untestable		96.5
	Untestable SFIT		
7	Spoiled	47	0.12
8	Unused	0	0
9	Expired	1311	3.39
10	Incomplete	0	0
11	Non-technical fail	36	0.10
12	Unresolved ID query	11	0.03
13	Technical Fail	0	0
14	Total Untestable	1405	3.5
15	**Tested and Untestable**	40125	
16	Response rate		60.6
17	Unlabelled	172	0.43
18	Undated	134	0.33

6. Return of Repeat samples

Table 4. Return of Repeat samples data for the complete study

Expired SFIT			Non Technical Fail SFIT			Spoiled SFIT		
Condition	**Number**	**%**	**Condition**	**Number**	**%**	**Condition**	**Number**	**%**
Returned	1134	86.5	Returned	34	94.4	Returned	41	83.7
Not returned	171	13.0	Not returned	2	5.6	Not returned	8	16.3
Total	1305		Total	36		Total	49	
Neg	1051	80.2	Neg	34	94.4	Neg	31	63.3
Pos	32	2.4	Pos	0	0.0	Pos	3	6.1
Expired	40	3.1	Expired	0	0.0	Expired	1	2.0
Spoiled	4	0.3	Spoiled	0	0.0	Spoiled	6	12.2
Closed	167	12.7	Closed	2	5.6	Closed	8	16.3
FOBT returned	6	0.5	FOBT returned	0	0.0	FOBT returned	0	0.0
FOBT not Returned	3	0.2	FOBT not Returned	0	0.0	FOBT not Returned	0	0.0
Others	8	0.6	Others	0	0.0	Others	0	0.0
Total	1311		Total	36		Total	49	

The return rate for repeat samples was more than 80% in each category. This is higher than the uptake of the initial screening test. Experience in the Programme indicates that the return rate of qualitative FIT is approximately 95%, which again is higher than uptake for the initial screening test.

7. Analysis of quantitative results

Statistical evaluation of the data from the study shows that there is a difference in faecal haemoglobin concentrations between males and females. Females had a faecal haemoglobin concentration of 283 ng/mL at the 97.5% percentile (95% CI: 257 – 316) and males 519 ng/mL (95% CI: 468 – 575). Confidence intervals in the data are very large, of course, because of the nature of population studies. To obtain 2.0% positivity, the cut-off concentration for females would have to be approximately 388 ng/mL and 684 ng/mL for males.

There is a steady increase in faecal haemoglobin concentration with age in both sexes. The faecal haemoglobin that would be used to give ca. 2.0% positivity in a 50 year old female would be 270 ng/mL, whereas the cut-off that could be used for a 70 year old female would be 680 ng/mL. There is a similar pattern seen in males. A 50 year old male could be reported as positive at 390 ng/mL, while his 70 year old counterpart reported positive at 950 ng/mL.

Table 5. Laboratory age and gender data for complete study

Males

	Returned	Percentiles					
	n	2.5	25	75	95	97.5	Cut off to give 2 % positivity
AGE							
50 - 54	4075	0	0	7	104	281	390
55 - 59	4160	0	0	9	154	415	525
60 - 64	3489	0	0	12	185	520	670
65 - 69	3497	0	0	17	253	713	950
70 - 74	2837	0	0	21	347	737	950
TOTAL	**18058**						

Females

	Returned	Percentiles					
	n	2.5	25	75	95	97.5	
AGE							
50 - 54	4543	0	0	6	68	170	270
55 - 59	4730	0	0	7	92	243	350
60 - 64	4058	0	0	8	101	235	340
65 - 69	3985	0	0	10	129	317	430
70 - 74	3346	0	0	15	192	534	680
TOTAL	**20662**						

Summary

- The number of participants who completed their cycle was 58.5%. This is higher than that reported in the Programme KPI.
- 98% of SFIT kits were returned to the Laboratory within 10 days of sample collection.
- In the Laboratory, turnaround time for analysis and reporting of SFIT results at 2 days is greater than 95%.
- The SFIT positivity during the course of the evaluation was 2.4%; this is slightly higher than planned.
- The number of participants who were sent a second test kit (except expired kits) is low compared to the Programme data.
- Analysis of quantitative data compares favourably with the published reports of other groups.
- The adverse weather led to problems for the Programme as well as the SFIT evaluation and this must be discussed in detail elsewhere

PJM/CGF/JAS

30 September 2011

Fit as a First Line Test
Summary of Call Recall Aspects

Introduction of a quantitative Faecal Immunochemical Test (FIT) as a first line test pilot (FFLT) within a population screening programme required careful planning and preparation. Call recall aspects included; training staff, preparing Helpline information, designing new instruction leaflets and agreeing lines to take regarding implications of introducing a new test within an existing programme. Careful consideration had been given to setting the test outcomes to reflect the same level of positive referrals as with the standard test kit. It was important that people who were offered this test kit understood this was not a study and that they were not receiving a lesser service. The following text was agreed to be inserted into the standard invitation letter:

- *You have been issued with a collection tube to return your sample. It is different from the kit you may have seen in other bowel screening leaflets. This is easier to use, offers the same high standard of result and can be tested by machine.*

The Know the Facts leaflet issued with the pack showed photographs of the standard test kit and it was hoped this sentence would reduce confusion over the different test kits in use at that time.

Helpline Information
Helpline staff were very involved in preparing the mail for these new kits as it is part of their role to prepare all mail for the Centre. They were familiar with the method of obtaining samples and trained in the expected method of sampling via the tube test. There were question and answer sessions held for staff and there was confidence that any calls would not be over complex.

Helpline Outcomes
There were very, very few calls with questions about how to do this test. 205 calls over a 6 month period of the 66,225 invitations issued. (0.3% of the invitations as compared to 6% currently). It was quickly obvious that participants could easily follow the new instructions. There was virtually no impact on the Helpline attributable to this test. There was some evidence that participants preferred the tube test and the single sample method to the standard kit. This evidence came in the form of enquiry to the Helpline for a tube test to be issued rather than the standard FOBT. Awareness of the availability of the new tube test appeared to be where two people in the same household were invited at different times and each had a different test kit due to the time period of the pilot.
See detailed information in Appendices 1 and 2

Information Materials
Feedback from a focus group who were known to have manual dexterity and/or visual impairment and the experience of call recall staff who listen to daily feedback from the public was used to assist in planning the instructions for doing the test.
Evidence from previous feedback was that opening the standard bowel screening invitation pack envelope could often land in the contents falling to the floor and the recipient unsure of what order the items should be read. The Welsh Screening programme utilises a paper wallet to avoid this problem. Effort was

taken to create a wallet to contain the items and this wallet was used to display all messages in a simple layout. The layout gave room for large font text and included a simple " IKEA" style format which had no text. It was expected that this would assist those who are unable to read English.
The 3 test items (the tube, the zip-lock bag and the return plasticised envelope) were all enclosed in a plastic flow pack which assisted hand filling of the wallets by staff at the Centre. This helped ensure components were complete, and gave the pack a more professional finished look.

The new wallet utilised:

- Clear arrows to point out the direction of the step by step instructions to rather than numbering.
- very accurate diagrams of the bar-coded tube test
- very clear illustration of the bar-coded label being removed from the letter and applied to the zip-lock bag.
- Exact diagrammatic replication of the bar code label
- replication of the checklist on the back of the return envelope
- for the first time an illustration of a roll of toilet paper
- for the first time an illustration of a bowel motion in the toilet
- a summarised view of the instructions without any text on the reverse

Information Materials Outcomes
There was valuable feedback previously obtained via a group known to have manual dexterity and/or visual impairments. This feedback was carried out using the standard Faecal Occult Blood test kit (FOBT) and was known as phase one. The same group were asked to feedback on the new FFLT pack and instructions and the key points are summarised as follow:

95% of participants found the print size to be accessible. This is compared to 85.3% of the volunteers in phase one where the standard FOBT kit and instructions were focus group tested.

- 100% of participants thought that the instructions were easy to follow. Many commented that the instructions were a great improvement on the original.
- Only 39.1% of participants encountered handling problems due to their disability using the tube collection device kit. This is compared to 61.8% of the volunteers in phase one.
- 78.3% of participants felt that the tube collection device kit was easier to use than the original.

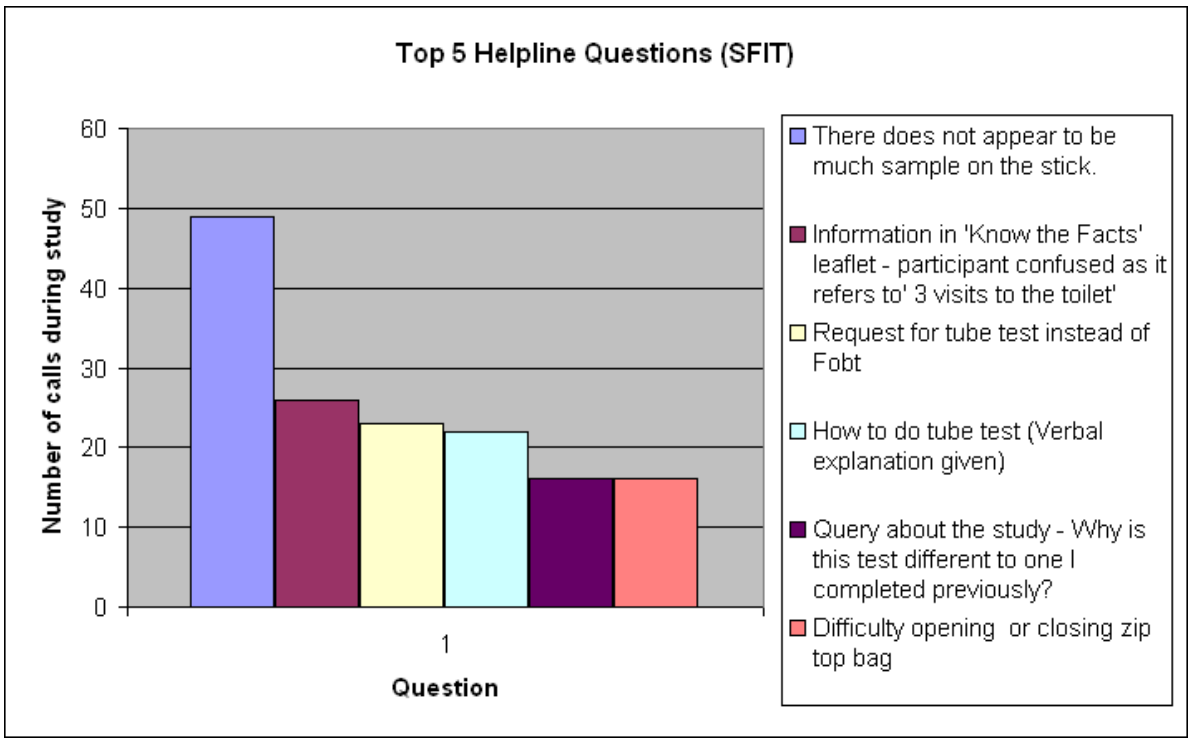
The wallet required to be hand inserted into an envelope with a folded letter and the Know the Facts Leaflet. This manual process would not be transferable to a national screening programme and would require full automation.

Conclusions
There were no issues encountered or reported by the public in the introduction of the FIT as a First Line Test pilot within a population screening programme in Scotland. Participants appeared to be able to follow the instructions without needing to call the Helpline for more information. A process of automation would need to be identified to prepare mail that contained and FIT tube test kit. The participants appeared to prefer the tube test and the single sample method to the standard kit.

Appendix 1

Top 5 Questions

There does not appear to be much sample on the stick.	49
Information in 'Know the Facts' leaflet - participant confused as it refers to' 3 visits to the toilet'	26
Request for tube test instead of FOBt	23
How to do tube test (Verbal explanation given)	22
Query about the study - Why is this test different to one I completed previously?	16
Difficulty opening or closing zip top bag	16



Appendix 2

Total Number of SFIT Invitation Packs issued 1st July 2010 - January 2011= 66,225 **Total number of Helpline calls received 1st July - 7th February = 31,070**	**Number of calls**
Only 205 calls specific to the SFIT kit were received in this 6 month period from 66,225 invited	
Helpline Questions during SFIT study (Data collected 1st July 2010 - 7th February 2011)	
Query about the study - Why is this test different to one I completed previously?	16
There does not appear to be much sample on the stick.	49
Query regarding expiry time	12
How to do tube test (Verbal explanation given)	22
Damage to tube by Child/ Animal/ other - concern about liquid content of tube	4
Should I not have 3 tubes to complete?	6
Difficulty opening or closing zip-lock bag	16
Are 3 samples not better than 1? (concern about intermittent bleeding)	7
I forgot to put tube into zip lock bag (label applied to tube or envelope)	4
I applied label to tube instead of bag	3
Information in 'Know the Facts' leaflet - participant confused as it refers to' 3 visits to the toilet'	26
Request for FOBt instead of tube test	2
Request for tube test instead of FOBt	23
Difficulty removing lid from tube	1
Participant thought zip lock bag was the envelope to return the kit	2
Concern that the label goes onto the bag instead of the tube (How will they know the sample is mine if the bag is separated from the tube?)	4
stick has broken from lid of tube	1
What should I do with the liquid in the tube?	2
Should I remove the liquid from the tube?	3

Should I fill the bag with stool sample?	1
Concern that excess sample seeped down side of tube	1
Total number of calls	**205**

Additional Comments made to helpline staff regarding the Tube Test

Participant thought test was fantastic - more hygenic, quick & easy

One participant accidentally stood on tube kit which burst open
One lady was angry that her tube test had expired by the time it reached the laboratory during extreme adverse weather - She felt that kits should not have been issued during periods of mail delays

Participant commented that he thought that the instructions were very clear
Gentleman stated that he had not completed Fobt kit 2 years ago as he believed it to be too 'fiddly' to complete, however he was more than happy to complete tube test as it seemed much easier.

One lady had requested a tube test for her husband as they both thought it was much 'easier' to complete than the Fobt kit

Participant commented that the tube test was a better, easier test to complete
One participant commented that they found the test much easier to do than the Fobt
One disabled person found the kit easier to use and felt it should be available for all disabled people.

One gentleman found the test to be a better test for him as due to his working hours, he found it difficult to collect 3 samples in 10 days.

One lady thought it was a more hygenic test as she did not like keeping the Fobt kit in the house between samples when she completed the test 2 years ago

3 participants commented that they preferred the Tube test as they found it easier to complete than the red and white card test

APPENDIX 5 SCOTTISH BOWEL SCREENING PROGRAMME INVITATION LETTER

Bowel Screening Centre
Kings Cross
Clepington Road
Dundee
DD3 8EA

NHS SCOTLAND

PRIVATE AND CONFIDENTIAL
A.N. OTHER
1 THE STREET
THE TOWN
XX1 1XX

Date 22 September 2015
Reference No: 0123456789

Enquiries to 0800 0121 833

Dear A.N OTHER

This is your invitation for Bowel Screening, which you are due to receive once every two years after the age of 50. Your test pack is enclosed. Please do not ignore this letter.

NHS Scotland offers bowel screening tests to all people aged between 50 and 74 every two years. The NHS register shows that you are within the age range to do the test.

Screening aims to find cancer of the bowel at an early stage when treatment is more simple and effective. The screening test looks for hidden blood in the bowel motion. This may be caused by a number of conditions, including bowel cancer. Doing the screening test is a good way of looking after your health. It's quick and easy and you do it all at home. The result comes back to you within 2 weeks.
Please follow the instructions carefully. Remember that you can call the free and confidential Helpline on 0800 0121 833 if you have any questions.
We will send you a reminder in 6 weeks. If you decide not to do the test you don't need to take any action. Your decision will not affect any health care you receive from your GP or hospital.

It's important that:

- You understand why you are doing the test
- You label your kit. You must use the label at the bottom of this page. Unlabelled kits will not be tested.
- You do the test as soon as you can
- Post it back as soon as possible in the freepost envelope

Yours sincerely

Professor Bob Steele
Clinical Director

****remove your label****

A.N. OTHER
0123456789
0011223344

IMPORTANT!

Peel this label off
And stick it on your
kit test

Helpline
0800 0121 833
Open 8 am to 5pm
Monday to Friday.
Answer phone available.
For Textphone users
(18001) 0800 0121 833

REGULAR SCREENING FOR BOWEL CANCER CAN SAVE LIVES
More information can be found at: www.bowelscreening.scot.nhs.uk

Bowel Screening:
Scottish Bowel Screening Programme

Scottish Bowel Screening Centre
Kings Cross
Clepington Road
Dundee
DD3 8EA

NHS SCOTLAND

PRIVATE AND CONFIDENTIAL

Date 24 September 2015
Reference No:

Enquiries to 0800 0121 833

Dear

You recently sent us your completed bowel screening kit for testing. Thank you for taking the time to do the test and post it to the Bowel Screening Centre.

Your test result shows that there were hidden traces of blood in your bowel motion (poo). This doesn't mean that you have cancer, but it does mean you should get the cause of the bleeding checked out.

There can be lots of simple explanations, including bleeding gums, inflammation in the stomach, polyps (non-cancerous growths) in the bowel, piles (haemorrhoids), or broken skin around your bottom (back passage). A more serious explanation could be that early bowel cancer sometimes bleeds.

A colonoscopy is the best way of looking for the cause of bleeding, which in some, but not all, cases may be due to bowel cancer.

What is a colonoscopy?

- It's usually an outpatient appointment, so you shouldn't need to stay in hospital for more than a few hours.
- A thin, flexible tube with a camera will be used to examine your bowel. This means the doctor or nurse can fully examine your bowel.
- The tube will reach your bowel by passing through your bottom (back passage).
- Your colonoscopy will take about half an hour and is a very safe examination.
- If 'polyps' are found, most of them can be removed at the time of the test. Polyps are small growths of cells on the bowel wall. Removing polyps can give you long-term protection from bowel cancer.

Remember, very few of the people who need a colonoscopy will have bowel cancer.

Your NHS Board will contact you in the next few days to explain the colonoscopy more fully to you. You will also be able to ask any other questions you may have about bowel screening.

If you wish to contact them yourself, the contact details are: …INSERT BOARD DETAILS HERE…

Your GP has a copy of your test result and will know that a colonoscopy is planned for you.

If you are still aged between 50 and 74 then we will invite you to take another bowel screening test in two years' time.

Yours sincerely

Professor Bob Steele
Clinical Director

Caron Innes
Research Nurse
Ninewells Hospital
Dundee

DD1 9SY
Name of Patient
Address

Date

Dear Sir/Madam,

Thank you for agreeing to take part in this Study.

I enclose an information sheet which provides more details about the study and how to collect the sample. It also contains the sample tube, a zip-lock bag and a return envelope for you to send everything back. Returning the sample will be taken as consent to participate in the study. Please remember to date the tube when you have taken the sample and post it as soon as possible.

You could bring your sample to your Colonoscopy appointment and hand it to one of the Nurses, if you prefer.

Please feel free to contact me if you have any further questions.

Yours faithfully

Caron Innes

Phonexxxxxxxx

PATIENT INFORMATION SHEET

Quantitative Faecal Immunochemical Testing in Patients with Large Bowel Symptoms. (A Potential New Screening Test)

You are being invited to take part in a research study which will be used as part of an educational qualification. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

If you wish this information sheet in another language please contact the research team who will be happy to assist.

Thank you for reading this.

1. What is the purpose of the study?

We hope to develop a stool test that can be used in the screening for some bowel disorders. This would be useful as it would hopefully reduce the number of people being referred for colonoscopy, as the test would identify people who did not need colonoscopy but another form of test.

2. Why have I been invited?

You have been invited because you are awaiting a colonoscopy, which your GP has requested, as you have some bowel symptoms. We would like to measure the amount of blood you have in your stool and compare it with the results of your colonoscopy.

3. Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep. Returning the sample to the research team will be taken as consent to participate in this study.

If you decide to take part you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time or a decision not to take part will not affect the standard of care that you receive.

4. What will happen to me if I take part?

You will be asked to kindly send us 1 sample from a bowel motion, **before** you start taking the medicine to clear your gut for the colonoscopy.

Your doctor will contact you with the results of your colonoscopy as usual.

5. What do I have to do?

We will give you 1 small tube and instructions that will enable you to take the stool sample. We will also give you a suitable stamped addressed envelope in which to return the sample to us.

To take the sample;

- Place layers of toilet paper in the toilet bowl to catch your sample. Do your stool sample onto the toilet paper. It doesn't matter if your sample touches the toilet water.
- Twist open the green cap of the tube and pull it out.
- Collect a sample by scraping the green stick on the bowel motion until the grooved end of the stick is covered.
- Place the stick back into the tube and push closed. Wash your hands.
- Put the date on the tube. Then place the tube into the zip-lock bag and close it.
- Put it in the return mailing envelope for posting.
- Review the checklist on the back of the envelope and post without delay.

6. What are the possible disadvantages and risks of taking part?

There are no risks involved for you taking part in this study. We understand that not everyone will feel able to take a stool sample from their own bowel movements and if this applies to you then we would understand why you would decline to participate in the study.

7. What are the possible benefits of taking part?

There is no benefit for you taking part in this research. We hope to use your donated samples to develop a test that will help people in the future.

8. Will my taking part in this study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential. All samples will be made anonymous with a unique study number to identify them, they will not be identified with your name.

We will use these samples in our current study. You will not be identified in this research. After the study all the samples will be destroyed.

9. What will happen to the results of the research study?

The results will be published in scientific journals and presented at meetings of scientists and doctors.

10. Who is organising and funding the research?

The study has been organised by Professor Fraser and his research team in the Scottish Bowel Screening Centre. The equipment is being provided by MAST diagnostics (the company which makes the test).

The Tayside Committee on Medical Research Ethics, which has responsibility for scrutinising proposals for medical research on humans, has examined this proposal and has raised no objections from the point of view of medical ethics. It is a requirement that the research records are made available to monitors from NHS Tayside, whose role is to check that research is properly conducted and the interests of those taking part are adequately protected.

If you believe that you have been harmed in any way be taking part in this study, you have the right to pursue a complaint and seek any resulting compensation through MAST diagnostics who are acting as the research sponsor. Details about this are available from the research team.

11. Contact for further information

If you have any questions or you wish to obtain further information about this study, you may contact Caron Innes at;

Ninewells Hospital and Medical School
Dundee
Telephone numberxxxxxxxxxxxx

Thank you for considering taking part in this study.